Learning and Action in Uncertain Environments

Louise Marshall

A dissertation submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy

UCL

July 2017

I, Louise Marshall, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

_____

Louise Marshall,

July 2017

Abstract

Successful interaction with the environment requires flexible updating of our beliefs about the world. By learning to estimate the likelihood of future events, it is possible to prepare appropriate actions in advance and execute fast, accurate motor responses. According to theoretical proposals, humans track the variability arising from dynamic environments by computing various forms of uncertainty. Several neuromodulators have been linked to uncertainty signalling but comprehensive empirical characterisation of their roles in perceptual belief updating and motor response modulation has been lacking. This thesis interrogates the contributions of noradrenaline, acetylcholine and dopamine to human learning and action within a unified computational framework of uncertainty.

First, I use pharmacological interventions to characterise the impact of noradrenergic, cholinergic and dopaminergic receptor antagonism on individual computations of uncertainty during a probabilistic serial reaction time task. I develop and employ a hierarchical Bayesian model to quantify human learning and action under three forms of uncertainty. I propose that noradrenaline influences learning of uncertain events arising from unexpected changes in the environment, while acetylcholine balances attribution of uncertainty to chance fluctuations within environmental contexts or to gross environmental violations following a contextual switch. In contrast, dopamine supports the use of uncertainty representations to engender fast, adaptive responses.

Second, I extend these results by focusing on the effects of natural inter-individual variations in dopaminergic function. Specifically, I employ the same task and model to assess individual learning and action under uncertainty as a function of *COMT* genotype.

Third, I focus on the role of noradrenaline. Uncertainty computations have been linked to changes in pupil diameter, and pupil dilation to noradrenergic neuronal activity in the locus coeruleus. Combining an auditory probabilistic learning task, pharmacological manipulations, pupillometry and computational modelling, I demonstrate that pupil diameter offers an indirect measure of dynamic noradrenergic computations of environmental uncertainty and volatility.

Acknowledgements

I have had the pleasure of learning from many remarkable people during my time at UCL. Sven Bestmann inspired me to pursue a PhD and he has been a better supervisor than I could have asked for. His mentorship has been a consistent source of guidance, encouragement and support. He has been generous with his time and accommodating of my interests. I am very grateful to have worked in the stimulating and collaborative research environment Sven nurtures. Importantly, Sven has taught me the value of discussing a research problem over a pint. He is, however, yet to convince me of the virtues of his film tastes. Perhaps there's still time.

The work in this thesis would not have been possible without the invaluable input of many talented scientists. I would like to thank Chris Mathys for his tuition, patience and computational expertise. I am grateful to Archy de Berker, Joe Galea, Klaas Enno Stephan and Peter Dayan for many fruitful scientific discussions. I also thank Graziella Quattrocchi, Simon Little and Diane Ruge for providing clinical supervision of the pharmacological experiments, and Paul Hammond and Amy Peters for offering the technical support that made the experiments possible.

I will look back at my years at Queen Square with great fondness because of the wonderful people I met there. I am thankful to Koko Hasan, Arthur Buijink, Joe Galea, Gionata Strigaro and Michelle Balaratnam for their many insights into neuroscience, academia and life. My experience working on the 4th floor has been a joy and that is down to Ella Clark, Svenja Espenhahn, Graziella Quattrocchi, Wolfgang Strube, Josh Hadwen and Andy Watson. Always there to share a cup of tea, they have smoothed any bumps in the winding road to completing a doctorate. I feel privileged to call you friends, and I look forward to many more travels, cultural anecdotes, and plank challenges.

I have been fortunate to live with some incredible housemates over the past four years. Chris – thank you for always being there to offer a listening ear, a glass of brandy, and a sprinkling of immoral support. Kiron – there is nobody else I would rather have ridden this PhD rollercoaster with; thank you for my culinary, televisual and musical education. Matt, Aodhín, Anna, Ollie and Chris D – thank you for making North London home.

Finally, I would like to thank my family for their continued love, generosity and good humour. Christine – I hope you enjoy this big book about the brain. Mum and Dad – I know we joke that I am accidentally turning into you, but I'd be hard pushed to find better role models.

Table of contents

List of figures

List of tables

Abbreviations

ACh	Acetylcholine
ACHE	Acetylcholinesterase
ACHE	Acetylcholinesterase gene
ACC	Anterior cingulate cortex
ADHD	Attentional deficit hyperactivity disorder
ALDH	Aldehyde dehydrogenase
ANOVA	Analysis of variance
BIS-11	Barratt impulsiveness scale
BPA	Bayesian parameter average
BOLD	Blood-oxygenation-level-dependent
BP	Blood pressure
CFQ	Cognitive failures questionnaire
COMT	Catechol-O-methyltransferase
COMT	Catechol-O-methyltransferase gene
CSE	Corticospinal excitability
CT	Control task
DA	Dopamine
DAT	Dopamine transporter
DAT1	Dopamine transporter gene
dlPFC	Dorsolateral prefrontal cortex
DNA	Deoxyribonucleic acid
DOSPERT	Domain-specific risk taking
DRD2	Dopamine D2-receptor gene
fMRI	Functional magnetic resonance imaging
FDR	False discovery rate
FOM	Forgetful Observer Model
GPCR	G protein-coupled receptor
HGF	Hierarchical Gaussian Filter
HR	Heart rate
IC	Inferior colliculus
INDEL	Insertion/deletion polymorphism
ITI	Intertrial interval
LC	Locus coeruleus
Met	Methionine

MRI	Magnetic resonance imaging
NA	Noradrenaline
NET	Noradrenaline (norepinephrine) transporter
NET	Noradrenaline transporter gene
ns	Non-significant
PCC	Posterior cingulate cortex
PCR	Polymerase chain reaction
PE	Prediction error
PET	Positron emission tomography
PFC	Prefrontal cortex
PLT	Probabilistic learning task
PPN	Pedunculopontine tegmental nucleus
PSRTT	Probabilistic serial reaction time task
RM-ANOVA	Repeated-measures analysis of variance
RT	Reaction time
RW	Rescorla-Wagner
SC	Superior colliculus
SD	Standard deviation
SED	Standard error of the difference
SEM	Standard error of the mean
SN	Substantia nigra
SNP	Single nucleotide polymorphism
SNRI	Selective noradrenaline reuptake inhibitor
SPECT	Single-photon emission computed tomography
TM	Transition matrix
TD	Temporal difference
Val	Valine
VAS	Visual analogue scales
VNTR	Variable number tandem repeat
VTA	Ventral tegmental area
YD	Yu and Dayan
9R	9-repeat allele
10R	10-repeat allele

1 Introduction

This thesis addresses the neuromodulatory mechanisms employed by the human brain to support learning and action in uncertain environments. It builds on a large body of theoretical, physiological, pharmacological, behavioural and computational work proposing roles for noradrenaline (NA), acetylcholine (ACh), and dopamine (DA) in computing different forms of uncertainty, and in supporting adaptive motor responses to environmental events. In this chapter, I introduce three theorised forms of environmental uncertainty, and review the existing literature on the neuromodulatory bases of uncertainty representations and response modulation. I highlight several open questions addressed in this thesis, define the key terms that will be used throughout, and present an overview of the following chapters.

1.1 Uncertainty is an inherent feature of the environment

Successful interaction with the environment requires flexible updating of our beliefs about the world (Conant and Ashby, 1970; Körding and Wolpert, 2004; Yu and Dayan, 2005; Behrens et al., 2007; O'Reilly, 2013). By tracking the environment's regularities, an individual can form and manipulate internal estimates of the world's statistical structure, and learn the causes of their sensory input. In so doing, it becomes possible to predict the likelihood of future environmental events given particular sensory cues (Friston, 2005; Bar, 2009), in turn facilitating anticipatory action preparation and the execution of fast, accurate motor responses (Bestmann et al., 2008).

However, the world with which humans, and indeed all animals, interact is incredibly complex; a multitude of statistical dependencies relate the sensory stimuli and events within our current environment, and these relationships are liable to change over time. Further, random events can occur due to environmental stochasticity. While our senses offer a means by which to track the myriad of entities within our environment, they only give us partial access to the true relationships that exist between entities. As such, the environment's richly complicated sources of noise and latent structure present us with various forms of uncertainty.

For instance, a London commuter predicting her journey time to work faces three distinct forms:

First, there is ***irreducible uncertainty***, which captures the randomness inherent in any complex environment and is undiminished by learning. An unplanned station closure or

a faulty train could cause journey delays and thus influence the accuracy of the commuter's estimated arrival time on any given day.

Second, ***estimation uncertainty*** arises from the commuter's incomplete knowledge of the probabilistic relationships *within* her current environmental context. After moving to a new part of town, the duration of the commuter's chosen route to work may be unclear, producing uncertainty about how likely she is to arrive at work on time. Over repeated journeys, this estimation uncertainty falls as the commuter learns the contextual rules of her environment. For example, she learns to predict the frequent delays on this new route due to congestion during the morning rush hour, although these delays may vary with local dips and surges in the number of passengers using the service.

Third, ***volatility uncertainty*** arises from the commuter's beliefs about the stability of the environment, and thus how quickly probabilistic relationships are changing *between* contexts. A major sporting event, such as the London Olympics, may bring a large influx of additional passengers for an unknown period of time and with unexpected effects on transport performance, making it harder to predict future journey times until these changes have been learned.

1.1.1 The brain computes different forms of uncertainty

To formulate accurate predictions about the likelihood of future events, and thus facilitate anticipatory preparation of appropriate motor responses, it is necessary to take these forms of uncertainty into account (Ma and Jazayeri, 2014; Meyniel et al., 2015; Pouget et al., 2016). In line with this notion, an assortment of theoretical, behavioural and neurobiological research has suggested that the brain computes uncertainty estimates relating to the environment's sensory events, contextual associations and their changes over time (Averbeck et al., 2006; Ma et al., 2006; Behrens et al., 2007; den Ouden et al., 2010; Fiser et al., 2010; Mathys et al., 2011, 2014; Payzan-LeNestour and Bossaerts, 2011; Bach and Dolan, 2012; Bland and Schaefer, 2012; Friston et al., 2012; Iglesias et al., 2013; Payzan-LeNestour et al., 2013; Vossel et al., 2014a, 2014b; de Berker et al., 2016; Diaconescu et al., 2017).

Uncertainty estimates influence our perceptual beliefs about the world. In psychology, a distinction has been made between two forms of information processing within the brain (Gregory, 1970, 1997). While bottom-up processing focuses on incoming sensory information from the environment, top-down processing uses past experience to guide the interpretation of environmental data in an expectation-driven manner. Uncertainty about the validity of one's own perceptual beliefs about the world should have the effect

of suppressing top-down prior expectations relative to new bottom-up sensory evidence, promoting learning about the sensory stimuli and events within the current environmental context (Yu and Dayan, 2003). With their broad distribution and extensive connectivity, the brain's neuromodulatory networks are well-placed to facilitate the widespread changes in neuronal gain required to support such a function (Berridge and Waterhouse, 2003; Aston-Jones and Cohen, 2005a; Warren et al., 2016). Indeed, neuromodulators profoundly alter the dynamics and topology of cortical networks (Marder, 2012; Eldar et al., 2013; Polack et al., 2013; McGinley et al., 2015). In particular, NA and ACh are known to enhance bottom-up, feedforward thalamocortical transmission of sensory information relative to top-down, intracortical and feedback processing (Hasselmo et al., 1996; Gil et al., 1997; Kimura et al., 1999; Kobayashi et al., 2000; Yu and Dayan, 2002, 2005; Hasselmo and McGaughy, 2004; Sarter et al., 2005; Dayan and Yu, 2006a; Deco and Thiele, 2011; Moran et al., 2013).

Figure 1.1 A schematic of the noradrenergic and cholinergic networks. *Both neuromodulatory systems show a broad distribution and extensive connectivity. (A) Noradrenaline (NA), also known as norepinephrine, is a catecholamine synthesised from an amino acid precursor, tyrosine, via a sequence of enzymatic steps (Cooper et al., 2003). The primary source of NA is a brainstem nucleus called the locus coeruleus (LC), which innervates the cortex, cerebellum and hippocampus (Sara, 2009). Functionally, NA has been linked to arousal and attention. (B) A major source of acetylcholine (ACh) is the basal forebrain, located below the striatum, which sends cholinergic projections to the cortex and hippocampus. An additional source of ACh lies within the pedunculopontine nucleus and laterodorsal tegmental nucleus of the brainstem. Functionally, ACh has been linked to arousal, attention and memory (Himmelheber et al., 2000; Jones, 2005). Figure adapted from Purves et al., 2011.*

1.1.2 Neuromodulatory computations of uncertainty

A seminal body of work by Yu and Dayan has had lasting impact on the theorised contributions of NA and ACh to uncertainty computations (Yu and Dayan, 2002, 2003, 2005; Dayan and Yu, 2006b). Specifically, the authors hypothesised that ACh signals the uncertainty that arises from ignorance about, and the unreliability of, a stable set of probabilistic relationships that link sensory events *within* an environmental context. As such, the quantity is notionally similar to estimation uncertainty. In contrast, Yu and Dayan suggest that NA signals the uncertainty that arises from unexpected events that occur *between* environmental contexts, i.e., following a contextual switch. Contextual switches arise due to environmental volatility and bring about a change in environmental rules. As I will address next, two types of experimental paradigm have highlighted different cholinergic and noradrenergic effects on behaviour within and between environmental contexts respectively, supporting a notional functional dichotomy for the two neuromodulators.

1.1.2.1 A proposed role for acetylcholine under estimation uncertainty

Within a stable environmental context, humans and animals show faster, more accurate responses to validly and predictably cued events than to those believed improbable (Posner, 1980; Downing, 1988; Bowman et al., 1993; Vossel et al., 2014b). This so-called validity effect is modulated by pharmacological (Witte et al., 1997; Phillips et al., 2000a), surgical (Voytko et al., 1994; Chiba et al., 1999), and neurodegenerative (Parasuraman et al., 1992) manipulations of ACh. Specifically, reaction times (RTs) to invalidly cued visual targets have been shown to decrease in both rats and rhesus monkeys following systemic injections of the cholinergic agonist nicotine, which boosts ACh neurotransmission. Similarly, RTs to invalidly cued targets are lower in human cigarette smokers compared to non-smokers (Witte et al., 1997; Phillips et al., 2000a). In each of these experiments, responses to validly cued targets were unchanged, meaning that the validity effect was either reduced or completely abolished. Conversely, the cholinergic (muscarinic) antagonist scopolamine, which reduces ACh neurotransmission, increases the validity effect in rats by disproportionately increasing RTs to invalidly cued targets (Phillips et al., 2000a). Moreover, lesions of the cholinergic basal forebrain in rats and monkeys have also been shown to selectively increase RTs following invalid cueing (Voytko et al., 1994; Chiba et al., 1999). The same behavioural effect is observed when patients with Alzheimer's disease, and thus a cholinergic deficit, are compared to age-matched healthy individuals (Parasuraman et al., 1992).

Together, these results suggest a role for ACh in learning within environmental contexts defined by particular probabilistic rules. More recently, it has been demonstrated that blood-oxygenation-level-dependent (BOLD) activity in the human cholinergic basal forebrain reflects an individual's estimation uncertainty about the probabilistic relationships linking environmental cues and outcomes, as quantified by a computational learning model (Iglesias et al., 2013). Moreover, pharmacological cholinergic stimulation under the drug galantamine has been proposed to increase the rate at which humans learn probabilistic relationships under estimation uncertainty (Vossel et al., 2014a), supporting the idea that ACh enhances learning accorded to stimuli with uncertain predictive consequences (Bucci et al., 1998) by suppressing the use of outdated top-down cues and boosting bottom-up sensory processing (Yu and Dayan, 2005).

1.1.2.2 A proposed role for noradrenaline under environmental volatility

While NA plays no consistent role in learning within environmental contexts (Clark et al., 1989; Witte and Marrocco, 1997), it is thought to offer an interrupt signal when volatility uncertainty arises *between* contexts (Clark et al., 1989; Arnsten and Contant, 1992; Smith et al., 1992; Coull et al., 1995; Witte and Marrocco, 1997; Bouret and Sara, 2005; Dayan and Yu, 2006b). Learning to make accurate predictions from the strongly unexpected observations that follow a contextual switch necessitates heightened sensory vigilance and a disregard for outdated top-down expectations. NA, with its role in regulating arousal and its broad neural network capable of triggering multiple, simultaneous changes across the brain (Bouret and Sara, 2004), is well-placed to rapidly coordinate this process.

At a cellular level, NA increases neuronal gain by boosting the efficacy of synaptic interactions between neurons and thus increasing the responsivity of target neurons to their afferent input (Servan-Schreiber et al., 1990; Berridge and Waterhouse, 2003; Aston-Jones and Cohen, 2005a; Warren et al., 2016). This noradrenergic effect on synaptic transmission within cortical structures is believed to upregulate the processing of external sensory stimuli relative to intrinsic top-down information (Hasselmo, 1995), therefore promoting experience-dependent neuronal plasticity (Harley, 1987; Sara et al., 1994; Aston-Jones et al., 1997; Bouret and Sara, 2005; Yu and Dayan, 2005; Corbetta et al., 2008; Tully and Bolshakov, 2010). By selectively increasing gain following unexpected sensory events that accompany a change in environmental context, the neuromodulator would be well-positioned to regulate an individual's learning rate under environmental volatility. Indeed, neurons in the locus coeruleus (LC), the primary source

of cortical NA, show strong responses to unexpected environmental changes in rats and non-human primates (Sara and Segal, 1991; Aston-Jones et al., 1997).

Pharmacological manipulations of NA have been shown to alter performance during tasks that feature contextual switches. For instance, administration of idaoxan, an α2-adrenoceptor antagonist that increases both the firing rate of LC neurons and noradrenergic release in the cortex and hippocampus, accelerates the detection of unexpected switches in the predictive properties of sensory stimuli in rats required to use visual or spatial cues to navigate a linear maze (Devauges and Sara, 1990). Similarly, systemic administration of an alternative α2-adrenoceptor antagonist, atipamezole, improves attentional set-shifting in rats; an effect that is blocked by microinjection of the α1-adrenoceptor antagonist benoxathian (which decreases NA neurotransmission) into the medial frontal cortex, an area homologous to the primate dorsolateral prefrontal cortex (dlPFC) (Lapiz and Morilak, 2006).

It has also been demonstrated that noradrenergic, but not cholinergic, deafferentation of rat medial frontal cortex impairs adaptation to contextual switches during attentional set-shifting tasks (McGaughy et al., 2008). Moreover, 6-hydroxydopamine-induced lesions of noradrenergic projections from the rat LC to the medial frontal cortex impairs set-shifting to novel stimuli (Tait et al., 2007). Importantly, since 6-hydroxydopamine can destroy both NA and DA neurons, the authors of this study verified that their neurotoxic lesioning method depleted NA in the medial frontal cortex, but caused no significant changes to DA neurotransmission. Further, systemic administration of atomoxetine, a selective NA reuptake inhibitor (SNRI) that increases extracellular NA concentrations, has been shown to improve attentional set-shifting in noradrenergically lesioned rats but has no effect in non-lesioned rats, highlighting the importance of optimal levels of cortical NA neurotransmission for optimal adaptive performance in dynamic environments (Newman et al., 2008).

With regards to whether humans depend on noradrenergic neurotransmission to detect and adapt to environmental volatility, BOLD activity in the human LC has been shown to dynamically track volatility uncertainty, as estimated by a computational learning model (Payzan-LeNestour et al., 2013). Moreover, pupil dilation, which is influenced by noradrenergic afferents (Joshi et al., 2016) correlates with unexpected events, such as those that occur due to changes in environmental context (Preuschoff et al., 2011; Nassar et al., 2012; Browning et al., 2015). Finally, SNRIs, thought to increase NA neurotransmission, are used to treat individuals with attentional deficit hyperactivity disorder (ADHD), a condition associated with deficits in reversal learning following

contextual switches and abnormal cortical catecholaminergic neurotransmission (Itami and Uno, 2002; Seu et al., 2009).

In sum, an extensive body of physiological, pharmacological, behavioural and neuroimaging work is compatible with the theory that ACh underlies learning of the relationships within stable environmental contexts, while NA supports learning under environmental volatility.

1.1.3 Motor responses are sensitive to uncertainty

As discussed above, representations of the uncertainty existing within and between environmental contexts are crucial for optimal predictions about the probability of future events. Optimal predictions, in turn, facilitate anticipatory preparation of appropriate motor responses. Indeed, previous work has demonstrated that the relationships between sensory events within probabilistic contexts can pre-emptively modulate the output of the motor system, thus speeding RTs to predictable events (Hick, 1952; Hyman, 1953; Requin and Granjon, 1969; Näätänen, 1970). Moreover, human corticospinal excitability (CSE), as measured with transcranial magnetic stimulation, has been shown to vary with uncertainty during a probabilistic RT task such that CSE is increased under low uncertainty about the required motor response to an upcoming event (Bestmann et al., 2008). Accordingly, high CSE is also accompanied by faster RTs.

However, good predictions are not in themselves sufficient for adaptive performance in dynamic environments. An additional mechanism is required to modify action selection based on one's own beliefs about the latent changes in the environment and/or the occurrence of unexpected events. Humans are indeed capable of engaging resources to inhibit a prepared response and replace it with an alternative when a unexpected event occurs (Hikosaka and Isoda, 2010; Isoda and Hikosaka, 2011), albeit at the expense of a prolonged RT (Galea et al., 2012; Bestmann et al., 2014). Unexpected events arise due to prediction errors that capture a mismatch between expectation and reality. As I will discuss in detail in **Chapter 3**, prediction errors provide the brain with an important teaching signal (den Ouden et al., 2012) that can trigger the modifcation of neuronal plasticity in target structures (Houk et al., 1995; Wickens et al., 2003; Frank, 2005), thus facilitating learning and behavioural flexibility.

1.1.3.1 A proposed role for dopamine in response modulation

DA neurons are known to fire in response to prediction errors (Mirenowicz and Schultz, 1994; Schultz et al., 1997; Schultz and Dickinson, 2000; Zink et al., 2003; O'Doherty et al., 2004; Bayer and Glimcher, 2005; Bunzeck and Düzel, 2006; Pessiglione et al., 2006;

Joshua et al., 2008; Matsumoto and Hikosaka, 2009; Zaghloul et al., 2009; den Ouden et al., 2010, 2010). Furthermore, there is considerable evidence linking DA to flexible behaviour (Figure 1.2). For instance, DA depletions due to Parkinson's disease are associated with specific flexibility impairments in both motor (Cools et al., 1984; Galea et al., 2012) and cognitive domains (Beatty and Monson, 1990; Cools et al., 2001a), with performance restored by dopaminergic medication (Cools et al., 2001b; Galea et al., 2012). Specifically, Parkinson's disease patients off dopaminergic medication show impaired switching between different cognitive task demands, such as naming letters or digits (Cools et al., 2001b). This effect, which is ameliorated by pharmacological DA stimulation, has been shown to be independent of both rule learning and working memory load since it occurs even when a contextual cue explicitly signals the required behaviour and any task switches (Cools et al., 2001a). Moreover, patients with Parkinson's disease have been shown to produce fewer motor responses in a finger-tapping task following a switch in the required finger-tapping sequence (Beatty and Monson, 1990). However, it should be noted that, in the latter study, patients' medication regimens were unchanged during testing sessions, meaning that pharmacological DA stimulation cannot be excluded as a possible confounding factor in this case.

Figure 1.2 A schematic of the dopaminergic network. *Dopamine (DA) is a catecholamine synthesised from the same amino acid precursor as NA, namely tyrosine. Like NA and ACh, DA neurons show a broad distribution and extensive connectivity in the brain. The principal sources of DA are the substantia nigra (SN) and the ventral tegmental area (VTA), both of which are components of the basal ganglia, located at the base of the forebrain. The SN sends dopaminergic projections to the dorsal striatum. This so-called nigrostriatal (or mesostriatal) pathway has been linked to motor, reward and associative learning functions. Dopaminergic neurons of the VTA project primarily*

to the prefrontal cortex via a mesocortical pathway, which has been linked to cognitive control and behaviour. A smaller group of DA neurons project from the VTA to the nucleus accumbens via a mesolimbic pathway linked to reward, aversion, pleasure and reinforcement learning. Together, the mesocortical and mesolimbic pathways constitute the mesolimbocortical pathway (Björklund and Dunnett, 2007). Figure adapted from Purves et al., 2011.

Recently, there has been renewed focus on the role of DA in modulating flexible motor responses. For example, Parkinson's disease patients off dopaminergic medication show an impaired ability to make adaptive responses to unexpected sensory events occurring within a broadly predictable context (Galea et al., 2012). Specifically, in a probabilistic serial RT task, responses to unexpected imperative stimuli, which elicit large sensory prediction errors and require replacement of a prepared action with an unprepared one, are slower than those made by healthy controls or by patients receiving dopaminergic medication. Importantly, the same effect is also observed when healthy individuals undertake the same task after having been administered the D1/D2-receptor antagonist haloperidol, which reduces DA neurotransmission (Bestmann et al., 2014).

Overall, it appears that DA plays a key role in modulating behavioural responses to low-level sensory prediction errors that necessitate motor flexibility. However, it remains unclear whether DA supports accurate response selection by facilitating perceptual belief updating (i.e., learning) in light of sensory prediction errors (Iglesias et al., 2013), or by modulating the sensitivity of motor response selection to perceptual beliefs.

1.2 A unified framework of uncertainty

To summarise, a considerable body of physiological, pharmacological, behavioural and theoretical work has suggested separable neuromodulatory involvement in the computations of, and responses to, uncertainty. However, attempts to characterise the relative contributions of NA, ACh and DA within a single computational scheme have been lacking. Computational models offer a sophisticated means by which to probe the brain's mechanisms of learning and action in uncertain environments. As such, they have brought significant advances to cognitive neuroscience in recent years (Daw et al., 2011; Takahashi et al., 2011; Iglesias et al., 2013; Diaconescu et al., 2017). By designing task paradigms in which key learning parameters change over time, and correlating these parameters with fluctuations in neural activity, it has been possible to infer the types of computations that underlie learning and behaviour. Indeed, in an assortment of studies of human learning, measures of neural activity have been linked to computations of

perceptual quantities such as uncertainty, prediction error and volatility (Hampton et al., 2006; Behrens et al., 2007, 2008; D'Ardenne et al., 2008; Hare et al., 2008; den Ouden et al., 2009; Cooper et al., 2010; Daw et al., 2011; Klein-Flügge et al., 2011; Boorman et al., 2013; Iglesias et al., 2013; Payzan-LeNestour et al., 2013; Diaconescu et al., 2017).

Given the array of perceptual quantities supposedly tracked by the brain, it is important to construct unified computational frameworks of uncertainty, prediction error and volatility. In so doing, it becomes possible to probe the relative contributions of different neuromodulatory systems to computing these quantities, and to modulating learning and action as the quantities fluctuate.

In the work I present in this thesis, I employ a Hierarchical Gaussian Filter (HGF) model (Mathys et al., 2011, 2014), in conjunction with a series of probabilistic learning tasks, to capture individual human learning under three distinct forms of uncertainty:

1. ***Irreducible uncertainty*** arising from the randomness inherent in any probabilistic environment;

2. ***Estimation uncertainty*** arising from an individual's incomplete knowledge of the probabilistic rules underlying the current environmental context;

3. ***Volatility uncertainty*** arising from the instability of these probabilistic rules over time.

Further, I develop a novel instantiation of the HGF to track the modulation of motor responses that occurs in light of an individual's uncertainty estimates.

As I will discuss in detail in **Chapter 3**, the HGF was first introduced by Mathys et al. as a generic hierarchical Bayesian framework for individual learning under the various forms of uncertainty inherent in the environment (Mathys et al., 2011). It has been successfully applied in several recent studies of probabilistic learning under volatility (Iglesias et al., 2013; Diaconescu et al., 2014, 2017; Hauser et al., 2014; Vossel et al., 2014a, 2014b, 2015; de Berker et al., 2016). The core component of the HGF is a three-level *perceptual model* that tracks an individual's learning about the environment's underlying structure. A novel second component, which will be introduced formally in **Chapter 3** and applied in **Chapters 4** and **5**, is a *response model* that maps an individual's beliefs about the environment, as provided by the perceptual model, onto his/her observed behaviour, here RT responses. This extension of the HGF makes it is possible to estimate the degree to which an individual's perceptual beliefs influence his/her motor responses.

For this thesis, I sought to characterise the relative contributions of NA, ACh and DA to computations of distinct forms of environmental uncertainty. Further, I aimed to disentangle the effects of the neuromodulators on individual perceptual belief updating from any effects on the sensitivity of motor responses to perceptual estimates. In the following chapters, I utilise two probabilistic learning tasks and a unified computational framework of uncertainty to quantify individual learning and response modulation in dynamic, probabilistic environments. In a series of experiments, I aim to pinpoint the relative impact of NA, ACh and DA on learning and action by utilising:

1. Pharmacological manipulations of NA, ACh and DA;
2. Genetic characterisation of baseline DA function;
3. Pupillometric measures of dynamic NA neurotransmission.

1.3 Pharmacological manipulations of neuromodulatory function

The notion that endogenous neuromodulators and exogenous drugs produce their physiological effects by interacting with cellular receptors was first introduced by John Newport Langley and Paul Ehrlich at the beginning of the twentieth century (Cooper et al., 2003). The idea was based largely on observations that some drugs could trigger specific biological responses while others prevented them. Since then, advancements in electrophysiological and pharmacological brain slice techniques, and the development of molecular cloning (Caulfield, 1993; Gingrich and Caron, 1993; Schwinn et al., 1995), have facilitated the identification of a vast array of cellular receptors. Drugs that bind to these receptors offer a useful tool with which to modulate endogenous neuromodulatory function. By combining pharmacological manipulations with cognitive tasks, it is possible to identify the contributions of different neuromodulators to human learning and behaviour (Pessiglione et al., 2006; Stelzel et al., 2010; Beierholm et al., 2013; Bunzeck et al., 2013; Chowdhury et al., 2013; Galea et al., 2013; Bestmann et al., 2014; Guitart-Masip et al., 2014; van der Schaaf et al., 2014; Vossel et al., 2014a; Crockett et al., 2015; den Ouden et al., 2015; Rutledge et al., 2015; Tomassini et al., 2015; Jepma et al., 2016; Warren et al., 2016; Diederen et al., 2017).

1.3.1 Pharmacological modes of action

Drugs can induce biological effects in several ways. A pharmacological agonist has both affinity and efficacy for a receptor, meaning that it can bind to that receptor and produce the same biological response as the receptor's endogenous ligand. An antagonist has affinity, but no efficacy, for a receptor. As such, it attenuates or blocks the biological

response produced by the receptor's endogenous ligand by competing with the ligand for receptor binding sites. An inverse agonist binds to a receptor but triggers a biological response opposite to that of the endogenous ligand (Stephenson, 1997; Bradley, 2014).

1.3.2 Types of receptor

1.3.2.1 Ionotropic receptors

There are two main classes of membrane-localised receptors: ionotropic and metabotropic. Ionotropic receptors are ligand-gated ion channels. They are composed of multiple subunits and a central pore. When an ionotropic receptor is activated by an endogenous ligand or pharmacological agonist, the pore opens, permitting the passage of Na^+, K^+, Ca^{2+} or Cl^- ions. The change in ion permeability can trigger excitatory or inhibitory action. Specifically, the influx of positively charged cations evokes depolarisation of the membrane potential, while the influx of negatively charged anions evokes hyperpolarisation, in turn making action potential firing more or less likely, respectively. For instance, ACh acts as an endogenous ligand at ionotropic (nicotinic) ACh receptors, and evokes excitation by increasing membrane permeability to NA^+ and K^+ ions. The effects mediated by ionotropic receptors are fast, occurring within milliseconds (Cooper et al., 2003).

1.3.2.2 Metabotropic receptors

Metabotropic receptors are G protein-coupled receptors (GPCRs). They are activated when an endogenous ligand or pharmacological agent binds to the receptor, inducing a conformational change and triggering an intracellular signalling cascade that ultimately results in the phosphorylation (or dephosphorylation) of proteins, and therefore protein activation, inactivation or functional modification. As such, GPCRs mediate slower responses (across seconds to minutes) than ionotropic receptors. These responses are generally modulatory, enhancing or dampening a neuronal signal. There are several classes of G proteins, including G_s, G_i, G_q and G_{12}, each of which activate different signal transduction pathways (Cooper et al., 2003). NA, ACh and DA act as endogenous ligands at different GPCRs located on the membranes of postsynaptic neurons within the brain. By administering pharmacological agents that interact with particular GPCRs, it is possible to disrupt the neuromodulatory function of the NA, ACh and DA systems.

1.3.2.3 Autoreceptors

At least within cognitive neuroscience, it is the ionotropic and metabotropic receptors located within the membrane of postsynaptic neurons that receive particular research

focus. Importantly, pharmacological agents acting at postsynaptic receptors have the capacity to modulate the firing rate of the postsynaptic cell in response to neuromodulator release from a presynaptic neuron. However, an additional presynaptic mechanism can also regulate postsynaptic firing. Presynaptic autoreceptors are sensitive to neuromodulators released by the neuron on which they are located. When a neuromodulator is released by the presynaptic neuron, it will activate these autoreceptors in addition to the receptors on the postsynaptic cell. Such presynaptic activation often serves as part of a negative feedback loop in signal transduction, with the autoreceptors typically inhibiting further release or synthesis of the neuromodulator, in turn modulating postsynaptic firing rate (Stephenson, 1997; Cooper et al., 2003; Bradley, 2014).

1.3.3 Reuptake and degradation of neuromodulators

Once a neuromodulator has been released by a presynaptic neuron, diffused across the synaptic cleft and activated receptors on the postsynaptic cell membrane, its action is terminated by a mechanism of reuptake and metabolic degradation. For instance, NA is absorbed back into the presynaptic neuron via reuptake mediated primarily by the noradrenaline transporter (NET). Once back in the cytosol of the presynaptic cell, NA is broken down by the enzyme monoamine oxidase (MAO), or repackaged into vesicles for future release. Similarly, DA reuptake is primarily mediated by the dopamine transporter (DAT). Once in the cytosol, DA is broken down into inactive metabolites by a set of enzymes that act in sequence: MAO, catechol-O-methyltransferase (COMT) and aldehyde dehydrogenase (ALDH) (Eisenhofer et al., 2004). ACh is inactivated primarily by the action of the enzyme acetylcholinesterase (ACHE), which catalyses degradation of the neuromodulator (Cooper et al., 2003).

Importantly, these transporters and degradative enzymes offer an additional means by which to study the function of the brain's neuromodulatory systems. Indeed, the action of NET, DAT, MAO, COMT and ACHE can be modulated using an assortment of drugs that target these proteins.

1.3.4 Pharmacological manipulations of noradrenaline, acetylcholine and dopamine

An understanding of the molecular and cellular mechanisms by which receptors, transporters and degradative enzymes regulate neuromodulatory function reveals relevant targets for psychopharmacological investigations. Indeed, by administering drugs known to interact with particular receptors, transporters and degradative enzymes,

it is possible to up- or down-regulate noradrenergic, cholinergic and dopaminergic neurotransmission in humans. Assessing any drug-induced changes in performance under carefully designed experimental paradigms offers the potential to implicate the different neuromodulatory systems in human learning and action under uncertainty.

1.3.4.1 Pharmacological manipulation of noradrenaline

The physiological targets of NA are the (nor)adrenergic receptors, otherwise known as adrenoceptors. Adrenoceptors are a class of metabotropic receptors with several subtypes: α1, α2 and β. A high density of α1-adrenoceptors exists in the human neocortex (Zilles et al., 1993). Since α1-adrenoceptors are the targets of noradrenergic neurons projecting from the LC, they form a sensible pharmacological target for investigations of the noradrenergic contributions to learning and action under uncertainty. In **Chapters 4** and **6**, I utilise prazosin, a drug that antagonises noradrenergic neurotransmission via a mechanism of inverse agonism at α1-adrenoceptors (Zhu et al., 2000), to investigate the contribution of NA to learning and action under uncertainty.

An alternative means by which to pharmacologically manipulate noradrenergic neurotransmission is to target the neuromodulator's reuptake machinery. For instance, the selective NA reuptake inhibitor (SNRI) reboxetine blocks the action of NET, in turn reducing the rate of NA reuptake from the synaptic cleft and supposedly increasing extracellular concentrations of NA (Wong et al., 2000). It has therefore been proposed that the drug's net effect is to increase NA neurotransmission. In **Chapter 6**, I use both reboxetine and prazosin in order to characterise the respective impact of up- and down-regulated NA neurotransmission on learning in uncertain environments.

1.3.4.2 Pharmacological manipulation of acetylcholine

There are two major classes of cholinergic receptors: nicotinic and muscarinic. Nicotinic receptors are ionotropic, while muscarinic receptors are metabotropic. There are five subtypes of muscarinic receptors (M1-5). Muscarinic M1-receptors, abundant in the neocortex and the hippocampus, are a major target of cholinergic neurons projecting from the basal forebrain (Volpicelli and Levey, 2004; Abrams et al., 2006). In **Chapter 4**, I utilise the M1-receptor antagonist biperiden to characterise the impact of reduced cholinergic neurotransmission on learning and action under uncertainty.

1.3.5 Pharmacological manipulation of dopamine

The physiological effects of DA are mediated via a class of metabotropic DA receptors. There are five known subtypes: D1-5. Research efforts have focused primarily on D1- and D2-receptors. D1-receptors are the most abundant dopaminergic receptors in the central nervous system. The highest concentrations of both D1- and D2-receptors exist within the basal ganglia, particularly in the caudate nucleus and putamen. Aside from the basal ganglia, D1-receptors have a wide distribution in the neocortex, amygdala and hippocampus. In contrast, the most significant densities of D2-receptors outside the basal ganglia occur in the hippocampus (Palacios et al., 1988; Hall et al., 1994). In **Chapter 4**, I administer haloperidol, a drug that, at sufficient doses, blocks both D1- and D2-receptors. The net effect of haloperidol is thought to be antagonism of dopaminergic neurotransmission. As such, I assess the impact of reduced dopaminergic neurotransmission on learning and action in uncertain environments.

1.4 Genetic variations in neuromodulatory function

An alternative means by which to examine the neuromodulatory underpinnings of learning and action is to adopt a behavioural genetics approach (Frank et al., 2007, 2009; Tan et al., 2007a, 2007b; Green et al., 2008; Ullsperger, 2010; den Ouden et al., 2013; Doll et al., 2016). It has been suggested that a considerable proportion of inter-individual variance in cognitive function can be accounted for by genetic factors (Friedman et al., 2008). Moreover, genetics influence the degree to which cognitive processes are disrupted under pharmacological manipulations, and in neurological and psychiatric disorders (Kimberg et al., 1997; Mehta et al., 2004a; Roesch-Ely et al., 2005; Frank and O'Reilly, 2006; Cools et al., 2007b; Clatworthy et al., 2009, 2009). Exploiting inter-individual differences in the genes that regulate neuromodulatory function offers an opportunity to identify the neural mechanisms that contribute to human learning and action under uncertainty. Additionally, a behavioural genetics approach permits the effects of different neuromodulatory systems to be assessed within individuals and in a single experimental session.

1.4.1 Types of genetic variation

Genetic variation can arise in several ways. First, a single nucleotide polymorphism (SNP) is a variation in a single nucleotide that occurs at a specific position on the genome (Sachidanandam et al., 2001). Second, a variable number tandem repeat (VNTR) is a location on the genome where a short nucleotide sequence is repeated, with the number of repeats commonly varying between individuals. Third, an insertion/deletion (INDEL)

polymorphism arises when a specific nucleotide repeat is present (insertion) or absent (deletion) (Rodriguez-Murillo and Salem, 2013). Approximately ten million SNPs and several thousand VNTRs and INDELs contribute to the vast genetic variation within human populations (Frank and Fossella, 2011). Among them are several genetic polymorphisms that influence noradrenergic, cholinergic and dopaminergic neurotransmission by modulating the availability of target postsynaptic receptors or the activity of neuromodulatory reuptake and degradation mechanisms. As I discuss next, the inter-individual variations in NA, ACh and DA function induced by these particular polymorphisms offer a potential means by which to characterise the neuromodulatory contributions to human learning and action under uncertainty.

1.4.2 *COMT*

The Val^{158}Met SNP in the *COMT* gene is one of the most widely studied polymorphisms in the behavioural genetics literature. The *COMT* gene encodes catechol-O-methyltransferase, an enzyme that catalyses the degradation of catecholamines, including DA. The COMT enzyme plays a significant role in regulating DA levels in the brain, providing the primary mechanism of DA degradation in the prefrontal cortex (PFC) (Gogos et al., 1998; Akil et al., 2003; Tunbridge et al., 2004; Yavich et al., 2007) but having little to no effect on striatal DA (Gogos et al., 1998; Sesack et al., 1998; Matsumoto et al., 2003; Tunbridge et al., 2004; Meyer-Lindenberg et al., 2005; Slifstein et al., 2008). The Val^{158}Met polymorphism results from a missense mutation that causes a nucleotide substitution from guanine to adenine, and therefore an amino acid switch from valine (Val) to methionine (Met), at rs4680 (codon 158). The Met isoform has reduced thermostability at body temperature, resulting in a three- to four-fold decrease in COMT enzymatic activity relative to the ancestral Val isoform (Männistö and Kaakkola, 1999; Chen et al., 2004). As such, synaptic DA concentrations are thought to be higher in Met carriers, particularly in the PFC. In contrast, the Val allele is associated with higher COMT activity and lower synaptic DA availability. Since the Val and Met alleles are codominant, heterozygotes show intermediate levels of COMT activity, explaining the trimodal distribution of COMT activity (corresponding to Val/Val, Val/Met and Met/Met genotypes) observed in human populations (Floderus et al., 1981).

It appears that COMT activity levels have decreased during human evolution (Palmatier et al., 1999; Chen et al., 2004), suggesting that the Met allele may have a beneficial effect on PFC function (Egan et al., 2001). Indeed, the Val^{158}Met polymorphism has been found to influence cognitive processing in various tasks that depend on the PFC (Egan et al., 2001; Malhotra et al., 2002; Goldberg et al., 2003; Winterer and Goldman, 2003;

Foltynie et al., 2004; Blasi et al., 2005; Frias et al., 2005; Meyer-Lindenberg et al., 2005; Diamond, 2007; Frank et al., 2007, 2009; Tan et al., 2007b; Diaz-Asper et al., 2008; Solís-Ortiz et al., 2010; Dumontheil et al., 2011). Carriers of the Met allele tend to show superior executive function compared to Val/Val homozygotes (Egan et al., 2001; Malhotra et al., 2002; Goldberg et al., 2003; Frias et al., 2005; Frank et al., 2007; Diaz-Asper et al., 2008). In the context of DA's proposed role in behavioural flexibility, it is particularly interesting to note that Met carriers are frequently better at switching between the demands of different task rules. For instance, task-switching performance in the Wisconsin Card Sorting Task, a widely used test of executive function in which individuals are required to match cards according to criteria that switch without explicit warning, is higher in Met carriers than Val/Val homozygotes (Egan et al., 2001; Malhotra et al., 2002). Since the Met allele has been linked to increased DA neurotransmission, this echoes the aforementioned finding of impaired task switching due to DA depletions in Parkinson's disease, which can be ameliorated with pharmacological DA stimulation (Cools et al., 2001b). Similarly, it has been shown that Met/Met homozygotes are more likely than Val carriers to switch their behaviour in response to instances of negative feedback (Frank et al., 2007), adding further weight to the proposal that DA has an underlying role in behavioural flexibility.

Although some alternative studies have observed the opposite effect of *COMT* genotype on executive function (i.e., that Val/Val homozygotes show better executive function than Met carriers), these experiments have not required behavioural task-switching. For instance, Parkinson's disease patients with a Val/Val genotype were found to demonstrate improved working memory and planning ability in the Tower of London task (Foltynie et al., 2004). However, it should be noted that this effect is likely to have been confounded by disease state and concurrent intake of dopaminergic medication. Higher task-switching performance has also been observed in post-menopausal women with a Val/Val genotype (Solís-Ortiz et al., 2010), but this finding is confounded by the fact that DA levels decline across the adult lifespan (Kaasinen et al., 2000; Bäckman et al., 2006; Li et al., 2010), inducing changes in cognitive performance (Bäckman et al., 2000; Dreher et al., 2008; Eppinger et al., 2013).

Nevertheless, the COMT enzyme appears to play an important role in regulating cortical DA neurotransmission, with associated effects on behavioural flexibility. More specifically, it has been proposed that the impact of DA on PFC function adheres to an inverted-U dose-response curve (Tunbridge et al., 2006). As such, it is thought that PFC-mediated cognitive processes are optimal within a relatively narrow range of intermediary DA activity, with insufficient or excessive baseline DA neurotransmission having a

relatively deleterious effect (Williams and Goldman-Rakic, 1995; Goldman-Rakic, 1998). Consistent with this notion, pharmacological studies in animals (Granon et al., 2000) and healthy human individuals (Kimberg et al., 1997; Mattay et al., 2000; Mehta et al., 2000) have shown that the effects of dopaminergic agents depend on baseline PFC function and *COMT* genotype (Mattay et al., 2003; Farrell et al., 2012).

In **Chapter 5**, I exploit *COMT* genotype to probe any natural inter-individual differences in DA-mediated behavioural flexibility by assessing learning and response modulation under uncertainty as a function of genotypic variations in cortical DA neurotransmission.

1.4.3 *DAT1*

Flexible learning and behaviour in dynamic environments depends not only on the PFC, but also on subcortical activity in the basal ganglia, particularly the striatum (Kehagia et al., 2010). The PFC and striatum interact via multiple serial and parallel loops (Alexander et al., 1986; Haber et al., 2000), which are under neuromodulatory influence. Indeed, cortical DA activity, regulated by COMT, has been shown to modulate subcortical DA neurotransmission (Grace, 2000), supposedly by indirect cortical feedback (Tunbridge, 2010). Further, it has been demonstrated that prefrontal lesions in rats can cause secondary impairments in striatal DA neurotransmission (Pycock et al., 1980).

A key target for investigations of striatal DA neurotransmission is the *DAT1* gene. A VNTR in the 3' untranslated region of *DAT1* influences expression levels of the dopamine transporter (DAT) (Mill et al., 2002). Since DAT plays a pivotal role in synaptic DA clearance in the striatum (Lewis et al., 2001; Frank and Fossella, 2011), the polymorphism modulates striatal DA availability (Caron, 1996; Heinz et al., 1999). The VNTR commonly occurs as nine (9R) or ten (10R) repeats of a 40 base-pair sequence, although between three and eleven repeats are known to exist (Forbes et al., 2009). The exact functional consequences of the polymorphism on DA neurotransmission are currently unclear. *In vitro*, the *DAT1* polymorphism causes natural variation in the expression of DAT (Mill et al., 2002). However, while some positron emission tomography (PET) and single-photon emission computed tomography (SPECT) studies have indicated that the 9R allele is associated with increased striatal DAT expression (Jacobsen et al., 2000; van Dyck et al., 2005; van de Giessen et al., 2009; Spencer et al., 2013), and a putative decrease in DA neurotransmission (Wichmann and DeLong, 1996), others have identified greater DAT expression in 10R carriers (Heinz et al., 2000; Fuke et al., 2001; Mill et al., 2002; VanNess et al., 2005). It should also be noted that a recent meta-analysis reported that there is currently no evidence to support the

hypothesis that the VNTR in the *DAT1* gene is significantly associated with inter-individual differences in DAT availability in the human striatum (Costa et al., 2011).

Nevertheless, the *DAT1* polymorphism has been associated with specific behavioural effects (Forbes et al., 2009; Gizer et al., 2009; Franke et al., 2010; van Holstein et al., 2011). Despite having no established effect on task-switching performance in a paradigm based on the Wisconsin Card Sorting Test (Garcia-Garcia et al., 2010), *DAT1* genotype has been linked to variations in behaviour during (rewarded) probabilistic reversal learning (den Ouden et al., 2013). Specifically, 9R carriers were found to make more post-reversal perseverative errors than 10R/10R homozygotes. This suggests that 9R carriers are more reliant on their previous experience, being more likely to select a previously rewarded stimulus following a contextual switch. This finding is compatible with rodent conditioning studies demonstrating that increased DA levels augment responses to previously rewarded stimuli (Parkinson et al., 1999; Goto and Grace, 2005). A related observation is that dopaminergic medication impairs reversal learning in patients with Parkinson's disease (Cools et al., 2001b), possibly due to dysregulation of reward-related dopaminergic processing in the ventral striatum (Cools et al., 2007a). Pharmacological inhibition of DAT under methylphenidate evokes similar impairments in healthy individuals (Clatworthy et al., 2009). Furthermore, increased reward-related activity is observed in the ventromedial striatum of 9R carriers (Dreher et al., 2009; Aarts et al., 2010). In light of this evidence, it has been suggested that the 9R allele may be associated with increased striatal DA concentrations, increased reward sensitivity and decreased behavioural flexibility. Alternative work has shown that flexible responses to a previously non-rewarded stimulus are impaired under neurological DA depletions due to Parkinson's disease (Peterson et al., 2009).

Despite conflicting reports of the impact of *DAT1* genotype on DAT expression and dopaminergic neurotransmission, the polymorphism holds the potential to offer finer insight into the contribution of DA to learning and action under uncertainty. In particular, assessing any impact of *COMT* and *DAT1* genotypes on learning and response modulation in uncertain environments could uncover evidence to suggest that specific processes are linked to cortical or striatal neurotransmission, respectively. One might hypothesise that, by altering cortical DA neurotransmission, different *COMT* genotypes might introduce variations in the ability to adapt behaviour in light of a (non-rewarded) switch in task demands. While the ventral striatum is linked to reward-related learning, the dorsal striatum is associated with motor function (Cools et al., 1984; Purves et al., 2011). As such, it is possible that altered striatal DA neurotransmission under different *DAT1* genotypes also influences an individual's ability to modulate motor responses

under environmental volatility. However, the aforementioned literature suggests that any effect of *DAT1* genotype on flexible behaviour is more likely to be linked to rewarded reversal learning than to rule-based behavioural task-switching.

1.4.3.1 DRD2

Inter-individual variation in DA-mediated task-switching behaviour has also been observed under different *DRD2* genotypes. The *DRD2* gene encodes the dopaminergic D2-receptor. D2-receptor density has been linked to individual capacity for switching between different task demands (van Holstein et al., 2011). A known SNP in the *DRD2* gene, namely the ANKK1-Taq1A polymorphism at rs1800497, results in an amino acid substitution from glutamic acid to lysine at position 713 and gives rise to the A1 allele. A1 carriers have a 30-40% reduction in *DRD2* expression compared to homozygous carriers of the A2 allele (Thompson et al., 1997; Ritchie and Noble, 2003), this effect being most prominent in the striatum, but also affecting the PFC (Noble, 2003). Increased D2-receptor expression in A2/A2 homozygotes has been linked to inferior (non-rewarded) task-switching performance compared to A1 carriers, which manifests as increased RTs, increased cortical switching-related activity, and increased functional connectivity in corticostriatal circuits (Stelzel et al., 2010), indicative of an association between D2-receptor density and increased task-switching effort. In line with this finding, it has also been demonstrated that pharmacologically stimulating human D2-receptors with the agonist bromocriptine increases switching-related activity in both the striatum and the posterior lateral frontal cortex (Stelzel et al., 2013). In contrast, the agonist decreases activity in sensorimotor regions supporting motoric hand-switching activity under task switches, indicating that dopaminergic stimulation likely has varying influences on different types of flexibility (e.g., cognitive and motor) due to complex interactions across the DA network. However, it should be noted that impaired task-switching behaviour has also been observed under D2-receptor antagonism with sulpiride (Mehta et al., 2004b).

The finding of a link between D2-receptor expression and task-switching is in line with studies that have related increased striatal D2-receptor density in schizophrenia (Wong et al., 1986; Abi-Dargham et al., 2000) to increased cortical DA and to deficits in behavioural and cognitive flexibility (Thoma et al., 2007). The fact that different studies have observed that both increased and decreased cortical DA neurotransmission can impair cognitive and behavioural flexibility likely speaks to the aforementioned inverted-U relationship between DA levels and executive function (Tunbridge et al., 2006; Vijayraghavan et al., 2007). Related to this notion, van Holstein et al. have highlighted

the importance of optimal D2-receptor signalling for adaptive human behavioural flexibility. Specifically, the researchers demonstrated that D2-receptor stimulation by bromocriptine could improve cognitive flexibility, but only in individuals with low baseline DA levels, as reflected by the aforementioned VNTR polymorphism in the *DAT1* gene (van Holstein et al., 2011). The behavioural effect of bromocriptine was abolished by pre-treatment with the D2-receptor antagonist sulpiride, providing further evidence that the effect is mediated via D2-receptors.

In light of the variations in behavioural flexibility associated with different levels of D2-receptor expression, examining any impact of *DRD2* genotype on learning and action under uncertainty might offer further insight into an underlying dopaminergic mechanism, particularly when contrasted with any effects of the *COMT* and *DAT1* genotypes.

1.4.4 *NET*

The *NET* gene encodes the NA transporter, which functions to reuptake extracellular NA and thus modulates NA neurotransmission. Although a SNP occurs in the promoter region of the gene at rs2242446, it has been studied far less extensively than the polymorphisms discussed thus far. Variations in *NET* genotype have been found to correlate with conditions such as ADHD, depression and alcohol dependence (Huang et al., 2008; Zhao et al., 2013; Oh and Kim, 2016), and with sensitivity to antidepressants that target the NA transporter (Owens et al., 2008; Jeannotte et al., 2009; Sekine et al., 2010), but any specific impacts of the polymorphism on noradrenergic neurotransmission and behaviour are currently unclear. *NET* genotype offers a potential means by which to probe inter-individual differences in NA-mediated learning and action in uncertain environments. However, any results would be speculative until the impact of the *NET* polymorphism on NA neurotransmission has been better established.

1.4.5 *ACHE*

The *ACHE* gene encodes acetylcholinesterase, an enzyme that hydrolyses, and therefore inactivates, ACh. It has been shown that, in patients with Alzheimer's disease and therefore cholinergic impairment, an A/A genotype is associated with a better response to treatment with the acetylcholinesterase inhibitor rivastigmine (Scacchi et al., 2009). However, again, any specific effects of the polymorphism on cholinergic neurotransmission and behaviour are currently unclear. Nonetheless, variations in *ACHE* genotype might offer a means by which to speculatively study inter-individual differences in ACh-mediated learning and action under uncertainty.

In sum, these five genes are of potential physiological relevance to the noradrenergic, cholinergic and dopaminergic mechanisms underlying learning and action in uncertain environments. As such, I had originally planned to characterise the effects of *COMT*, *DAT1*, *DRD2*, *NET* and *ACHE* genotypes on human learning and response modulation under irreducible, estimation and volatility uncertainty. However, as I will discuss in detail in **Chapter 5**, due to unexpected methodological constraints, I focus instead on inter-individual differences in dopaminergic neurotransmission evoked by *COMT* genotype.

1.5 Pupil diameter as a proxy for dynamic noradrenergic uncertainty computations

For half a century, pupil dilation at constant luminance has been considered a marker of central arousal (Hess and Polt, 1964; Kahneman and Beatty, 1966; Bradshaw, 1967; Kahneman et al., 1967; Beatty, 1982). Inspired by recent proposals that pupil diameter might offer an indirect measure of noradrenergic neural activity in the locus coeruleus (LC) (Rajkowski et al., 1993; Phillips et al., 2000b; Aston-Jones and Cohen, 2005a; Murphy et al., 2014; Varazzani et al., 2015; Joshi et al., 2016), and that NA might modulate learning under volatility uncertainty (Yu and Dayan, 2005; Payzan-LeNestour et al., 2013; Marshall et al., 2016), researchers have started to probe whether transient changes in pupil diameter can be used as a proxy for physiological autonomic processes that occur during behavioural tasks (Siegle et al., 2003; Aston-Jones and Cohen, 2005a; Critchley, 2005; Satterthwaite et al., 2007; Einhäuser et al., 2008; Hupé et al., 2009; Einhauser et al., 2010; Gilzenrat et al., 2010; Privitera et al., 2010; Jepma and Nieuwenhuis, 2011; Preuschoff et al., 2011; Fiedler and Glöckner, 2012; Nassar et al., 2012; Wierda et al., 2012; Eldar et al., 2013; de Gee et al., 2014; Browning et al., 2015; de Berker et al., 2016; Korn et al., 2016; van den Brink et al., 2016; Urai et al., 2017). The sensitivity of the pupil to such processes means that pupillometry might offer a simple, non-invasive and cost-effective tool with which to measure individual noradrenergic computations of uncertainty, without the need for pharmacological interventions or behavioural genetics analyses.

1.5.1 Pupil diameter as an indirect measure of noradrenergic neurotransmission

As I will discuss next, there is converging evidence from electrophysiology (Rajkowski et al., 1993; Aston-Jones and Cohen, 2005a; Varazzani et al., 2015; Joshi et al., 2016), pharmacology (Phillips et al., 2000c) and human neuroimaging (Samuels and Szabadi,

2008; Murphy et al., 2014) to suggest a relationship between NA and pupil dilation under constant luminance.

1.5.1.1 Electrophysiological evidence of a link between noradrenaline and pupil diameter

In the last decade, the theory that changes in pupil diameter are directly related to fluctuations in noradrenergic neuronal activity in the LC has been the focus of a considerable body of research (Aston-Jones and Cohen, 2005a; Gilzenrat et al., 2010; Nieuwenhuis et al., 2010; Jepma and Nieuwenhuis, 2011; Eldar et al., 2013). The LC is a brainstem nucleus in the dorsolateral pons. As mentioned previously, it is the primary site of NA synthesis and the principal source of NA for the cerebral cortices, cerebellum and hippocampus (Moore and Bloom, 1979; Aston-Jones and Cohen, 2005a). The notion that the pupil might offer an indirect measure of LC-NA activity was largely inspired by an observation that the baseline firing rate of a single neuron recorded in monkey LC aligned closely with simultaneously recorded changes in pupil diameter (Rajkowski et al., 1993; Aston-Jones and Cohen, 2005a).

While neuronal activity in several brain regions, including the inferior colliculus (IC), superior colliculus (SC), anterior cingulate cortex (ACC) and posterior cingulate cortex (PCC), has been associated with changes to pupil size (Wang et al., 2012; Ebitz and Platt, 2015), the link between LC and pupil diameter seems most direct. In 2015, Varazzani et al. provided novel data to suggest that LC neurons are involved in mediating changes in pupil diameter by demonstrating that spiking activity in the noradrenergic LC, but not the dopaminergic substantia nigra (SN), is positively correlated with changes in pupil size during decision-making in rhesus monkeys (Varazzani et al., 2015).

Importantly, in the last year, Joshi et al. have offered the most convincing evidence of a causal relationship between LC-NA activity and event-driven changes in pupil diameter. By recording activity in the LC, as well as the IC, SC, ACC and PCC, of rhesus macaques, the authors established that, during passive fixation, fluctuations in the firing rate of many of the recorded neurons in these different regions correlated with fluctuations in pupil size. However, when the LC, IC and SC were electrically microstimulated, it was the LC stimulation that was found to consistently trigger transient increases in pupil size in a 250-700ms window following stimulation onset (Joshi et al., 2016).

Moreover, the NA-LC system also appears to show phasic activations during perceptual decision-making (Rajkowski et al., 2004; Aston-Jones and Cohen, 2005a; Bouret and

Sara, 2005; Sara, 2009), presumably triggered by feedback connections from the PFC (Aston-Jones and Cohen, 2005a; Dayan, 2012). It is therefore possible that pupillary responses to perceptual estimates, such as prediction error, uncertainty and volatility, reflect noradrenergic activity in the LC.

1.5.1.2 Pharmacological evidence of a link between noradrenaline and pupil diameter

Pharmacological evidence also suggests that NA modulates pupil diameter in humans. Specifically, α2-adrenoceptor agonists, such as clonidine, which decease the activity of central noradrenergic neurons, have been shown to decrease baseline pupil diameter and increase spontaneous pupillary fluctuations. In contrast, α2-adrenoceptor antagonists, such as yohimbine, which have the opposite effect on central NA, have been shown to have the opposite effect on pupils, i.e., an increase in baseline pupil diameter and decreased pupillary fluctuations (Phillips et al., 2000b).

1.5.1.3 Human neuroimaging evidence of a link between noradrenaline and pupil diameter

Until recently, the lack of reliable non-invasive measures of LC activity had limited investigations of the functioning LC-NA system in humans. In 2006, Sterpenich et al. used fMRI to correlate human pupillary dilation with neural activity during emotional memory retrieval. Activations in an area of the dorsal tegmentum of the ponto-mesencephalic region were identified, consistent with (but not definitively indicative of) LC activity (Sterpenich et al., 2006).

The decreased signal-to-noise ratio in the brainstem, resulting from the effects of cardiac pulsation and respiratory movement, means that it has been traditionally difficult to accurately locate the LC's small structure using fMRI (Astafiev et al., 2010; Payzan-LeNestour et al., 2013). However, Shibata et al. demonstrated that it is possible to image the neuromelanin (a by-product of monoamine synthesis) contained in noradrenergic neurons of the LC (Graham, 1979; German et al., 1988; Shibata et al., 2006). Furthermore, Keren et al. developed a probabilistic LC atlas using high resolution T1-turbo spin echo MRI (Keren et al., 2009), which some researchers have used to locate LC responses under volatility uncertainty (Payzan-LeNestour et al., 2013). Furthermore, by taking advantage of these two methodological advances, Murphy et al. were able to use simultaneous pupillometry and fMRI to generate empirical evidence of a relationship between pupil diameter and BOLD activity in the human LC (Murphy et al., 2014).

1.5.2 A proposed link between pupil diameter and perceptual beliefs

At the start of the decade, there was a move to integrate pupillometry into contemporary studies of human learning and behaviour. Motivated by proposals that different behavioural states might be mediated by two modes of LC-NA activity (Usher et al., 1999; Aston-Jones and Cohen, 2005a), and that pupil diameter might reflect these LC-NA activity profiles, researchers quantified pupil dilation during shifts between exploitation and exploration behaviours. Behavioural exploitation is defined as engagement with a particular task, while behavioural exploration is characterised by switches between tasks. According to adaptive gain theory, two LC modes promote exploitation and exploration by adaptively adjusting the responsivity of cortical neurons (Aston-Jones and Cohen, 2005a). A phasic mode, characterised by an intermediate level of baseline LC activity and large phasic increases in noradrenergic activity, produces selective increases in neuronal responsivity to task-related stimuli. The phasic release of NA temporarily increases the responsivity (i.e., gain) of target cortical neurons to their afferent input, thereby potentiating the processing of task-relevant stimuli (Servan-Schreiber et al., 1990; Doya, 2002; Berridge and Waterhouse, 2003) and optimising performance in the current task (i.e., exploitation). In contrast, a tonic mode, characterised by elevated baseline LC activity, tonic NA release and the absence of phasic responses, produces a more enduring and less discriminative increase in neuronal responsivity. Although this impairs performance within the current task, it facilitates the disengagement of attention from that task and the processing of other non-task-related stimuli and/or behaviours (i.e., exploration).

Accordingly, changes in pupil diameter have been detected under exploration and exploitation behaviours (Gilzenrat et al., 2010; Jepma and Nieuwenhuis, 2011). Gilzenrat et al. showed that baseline pupil diameter decreases gradually when individuals engage in a new task, while increases in baseline pupil diameter are associated with decreases in task utility and upcoming task disengagement and exploration (Gilzenrat et al., 2010). Similarly, in a four-armed bandit task in which individuals aimed to maximise reward by making choices between four slot machines whose mean payoffs changed gradually and independently over time, exploratory choices were preceded by a larger baseline pupil diameter than exploitative choices. Furthermore, individual changes in baseline pupil diameter were predictive of an individual's tendency to adopt exploitation behaviour (Jepma and Nieuwenhuis, 2011).

Although perceptual estimates and decision variables were present in these task paradigms, they were not the principal focus of the experiments. Nevertheless, this work

inspired an important methodological shift; subsequent studies focused on developing quantitative models to formally test the hypothesised association between human pupil dilation and the perceptual estimates underlying learning and behaviour in uncertain environments. As I will discuss next, a range of studies have investigated the modulation of pupil diameter by perceptual quantities such as uncertainty, prediction error (commonly conceptualised in the pupil literature as *surprise*), and volatility.

1.5.2.1 Evidence that pupil diameter is modulated by irreducible uncertainty and surprise

Preuschoff et al. sought to explicitly quantify the impact of distinct perceptual estimates on pupil diameter during an auditory gambling task (Preuschoff et al., 2011). On each trial, two cards were drawn in succession from a deck numbered 1-10. Before the auditory presentation of either card, participants were required to place a monetary bet on whether the first or the second card would have a higher value. The paradigm was designed to dissociate perceptual estimates of expected reward, irreducible uncertainty (which the authors conceptualised as risk), and surprise (here conceptualised as risk prediction error). The first card served as a probabilistic cue and the second card as the trial outcome. Irreducible uncertainty about the trial outcome captured the inherent randomness of the probabilistic relationship between the cue and outcome cards, and therefore the unreliability of the trial outcome. Following cue presentation, irreducible uncertainty showed an inverted-U relationship, with maximal irreducible uncertainty occurring when the value of the first card was 5 or 6, and minimal uncertainty when the first card was 1 or 10. Surprise was defined as high when expected reward prior to the second card was positive and the actual outcome was a loss, or when the expected reward was negative and the actual outcome was a win. Surprise was low when the expected reward had been positive and the outcome was a win, or the expected reward had been negative and the outcome was a loss.

Both irreducible uncertainty and surprise were found to have modulatory effects on pupil diameter (Figure 1.3). More precisely, pupil dilation occurring after the presentation of the first (cue) card was increased when there was low irreducible uncertainty about the relative value of the second (outcome) card (i.e., when the first card was 1 or 10, meaning the second card was certain to be higher or lower, respectively) compared to when there were medium (first card was 2, 3, 8 or 9) or high (first card was 4, 5, 6 or 7) levels of irreducible uncertainty about the value of the second card (Figure 1.3A). Pupil dilation occurring after the presentation of the second card was augmented under high surprise compared to low surprise about the reward outcome (Figure 1.3C).

Figure 1.3 Pupil dilation under irreducible uncertainty and surprise. *(A) Pupillary dilation occurring between the presentation of the first and second cards was modulated by irreducible uncertainty. The pupil showed greater dilation if the outcome was certain (low uncertainty; first card was 1 or 10) than if there was high irreducible uncertainty about the outcome of the second card (high uncertainty; first card was 4, 5, 6 or 7). Pupils showed an intermediary dilatory response on trials with medium levels of irreducible uncertainty (first card was 2, 3, 8 or 9). (B) Significance of the difference between high and low uncertainty trials as presented in A. The horizontal line denotes an expected false discovery rate (FDR) of 5%. Times of significant difference fall above this line. (C) Pupil dilation after presentation of the second card was increased under high compared to low surprise about reward outcome. (D) Significance of the difference between high and low surprise trials according to C. Notation is the same as in B. Data in A and C are mean ± SEM. Figure adapted from Preuschoff et al., 2011.*

Assuming a relationship between NA and pupil dilation, the authors took the finding that post-outcome surprise (i.e., risk prediction error) modulated pupil diameter as indicative of a similar noradrenergic role for uncertainty as DA has for reward, namely the encoding of error signals. However, as I will address next, the finding that pupil dilation was

augmented under low (rather than high) irreducible uncertainty conflicts with the findings of alternative studies.

In my work with de Berker et al., we demonstrated independent evidence to suggest that pupil diameter is modulated by both irreducible uncertainty and surprise (de Berker et al., 2016). In a volatile probabilistic learning task, participants were required to make predictions about binary aversive outcomes (electrical shock/no electrical shock) based on binary probabilistic visual cues (cue 1/cue 2). The true probabilistic relationship between cues and outcomes was fixed within a contextual block but changed discretely every 20-40 trials, with maximal irreducible uncertainty occurring under a 0.5/0.5 cue:outcome probability. The HGF model was applied to the behavioural data to quantify individuals' trial-wise estimates of uncertainty and surprise. Baseline pupil diameter (i.e., pupil diameter at the time immediately preceding cue onset) was found to increase with irreducible uncertainty about the current cue:outcome relationship (Figure 1.4A). In contrast to Preuschoff et al.'s finding, irreducible uncertainty was also shown to increase pupil diameter across the course of the trial (Figure 1.4B). An additional positive effect of surprise (here capturing sensory prediction error) occurred approximately two seconds after outcome presentation, mirroring Preuschoff et al.'s finding.

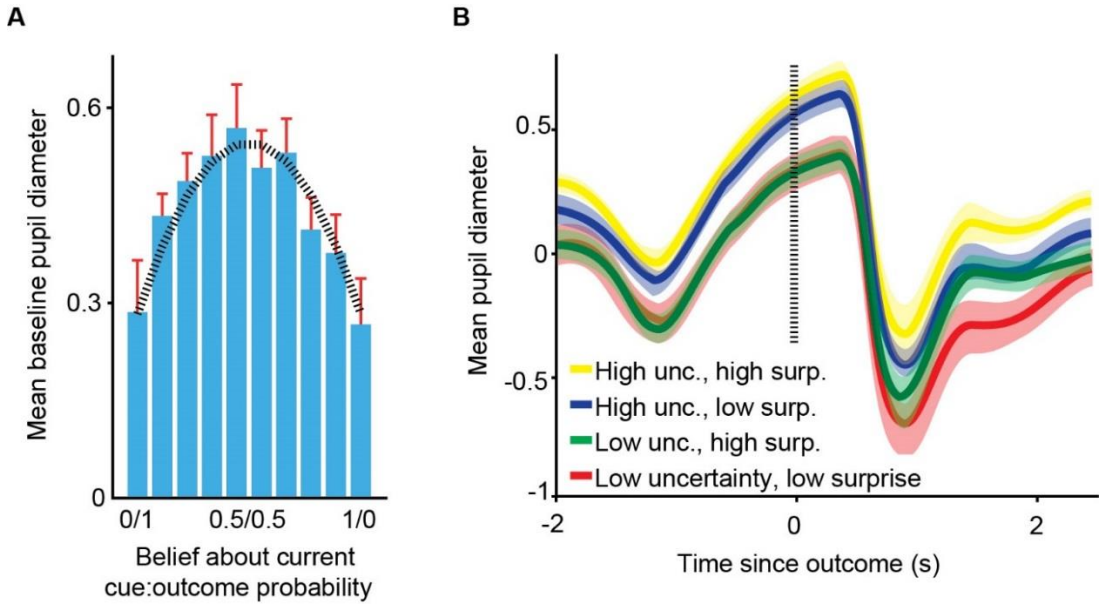

Figure 1.4 Pupil diameter is modulated by estimates of irreducible uncertainty and surprise. *(A) Baseline pupil diameter (i.e., pupil diameter immediately preceding cue onset) showed an inverted-U relationship with participants' beliefs about the current cue:outcome probability, as reflected by their irreducible uncertainty. The relationship closely conformed to a Bernoulli distribution (grey dashed line), with peak baseline pupil diameter coinciding with maximal irreducible uncertainty (i.e., when the estimated*

cue:outcome probability = 0.5/0.5). (B) Median splits, which separated trials according to whether they were high or low in irreducible uncertainty and high or low surprise, indicated that irreducible uncertainty increased pupil diameter throughout the trial. There was an additional positive effect of surprise approximately 2 seconds after outcome presentation. Data are mean ± SEM. Figure adapted from de Berker et al., 2016.

There are some important differences in the approaches adopted by Preuschoff and de Berker, which might explain why different effects of irreducible uncertainty on pupil diameter were observed. For example, the relative time-points at which the impact of irreducible uncertainty on pupil diameter was assessed differs between the two studies. While Preuschoff examined the pupillary effect of irreducible uncertainty after presentation of a probabilistic cue, de Berker probed any effect both at baseline (i.e., before cue onset) and across the trial time-course (relative to outcome presentation). There was no motivation to assess baseline pupil diameter in Preuschoff's study since pre-cue reward probability, and thus pre-cue irreducible uncertainty, were held constant (p=0.5). In contrast, in de Berker's task, the probabilistic cue:outcome relationship changed over time. Since irreducible uncertainty on the current trial could be computed on the basis of trial history (Mathys et al., 2011), it was possible to investigate whether baseline pupil diameter was modulated by this quantity.

Another difference between the two paradigms is that participants had to make predictions about different types of trial-wise outcomes (monetary reward/loss versus aversive/neutral stimuli), which may have evoked different neuromodulatory effects on pupil diameter. Further, these predictions were made at different time-points in the trial time-course. In de Berker's task, participants were required to predict the trial's outcome *after* the presentation of a probabilistic cue. In contrast, Preuschoff required participants to make a prediction about trial outcome *before* a probabilistic cue had been presented. As such, in Preuschoff's task, there are two additional parameters that could have augmented the post-cue pupillary response.

First, presentation of the cue would have conveyed a degree of post-decisional surprise. As mentioned previously, pre-cue irreducible uncertainty was constant in Presuchoff's paradigm as reward probability was fixed at 0.5. Depending on the direction of the participant's prediction and the value of the first (cue) card, trial-wise post-cue irreducible uncertainty would have increased or decreased at cue onset. For instance, if a participant had predicted that the second card would be lower than the first card, and the first card was then revealed to be a 2, they would have learned that there was a high likelihood that they had made a prediction error *and* that there was only moderate irreducible

uncertainty that the second card would be higher than the first. Post-decisional surprise would be highest when participants had made an incorrect prediction before the presentation of a cue card numbered 1 or 10 (i.e., on what Preuschoff et al. define as low irreducible uncertainty trials), possibly explaining the increased pupillary dilation on these trials. Second, given that low irreducible uncertainty in Preuschoff's framework actually reflected post-cue certainty about whether a participant would win or lose a monetary bet, the associated increase in pupillary dilation may, at least in part, reflect outcome confirmation. Indeed, monetary rewards and losses have been shown to have a positive effect on pupil diameter (Seymour et al., 2007).

In line with this suggestion, Satterthwaite et al. observed increased pupil dilation under *high* irreducible uncertainty when they utilised a similar version of Presuchoff's gambling task (Satterthwaite et al., 2007; Figure 1.5A). As in Preuschoff's paradigm, participants made predictions about the relative values of sequentially-presented pairs of cards, but here the prediction was made *after* presentation of the first (cue) card. This suggests that, when unconfounded by post-decisional surprise or outcome confirmation, irreducible uncertainty has a positive effect on pupil diameter.

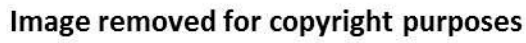

Figure 1.5 Pupil diameter tracks responses to uncertainty and surprise. *(A) Pupil diameter during both the post-cue and post-outcome periods was greater on trials with high irreducible uncertainty than on certain trials. This finding is in line with that of de Berker et al., but opposes the effect observed by Preuschoff et al. (B) Following outcome presentation, losses evoked increased pupillary dilation compared to wins on uncertain trials. (C) This effect was even larger on probable trials when the loss was relatively unexpected, i.e. when there was a larger prediction error that conveyed increased surprise. Data are mean ± SEM. Figure adapted from Satterthwaite et al., 2007.*

Moreover, pupillary dilation following presentation of the second (outcome) card, was augmented on trials on which participants had made a prediction error that led them to experience a monetary loss (Figure 1.5B). This effect was increased when the opposite outcome had been more likely (Figure 1.5C), echoing my hypothesised impact of post-decisional surprise on low irreducible uncertainty trials in Preuschoff's task. Nonetheless, it should be noted that, assuming participants made correct predictions on approximately 50% of trials in Preuschoff's task, any pupillary effect of post-decisional surprise may actually have been averaged out across trials.

1.5.2.2 Evidence that pupil diameter is modulated by volatility

Despite some differences in the precise results of these three studies, the observation that irreducible uncertainty and surprise modulate pupil diameter gives weight to the notion that pupil dilation offers an indirect measure of an individual's perceptual estimates. Given that noradrenergic neurotransmission has been linked both to changes in pupil diameter and to learning under environmental volatility, one might expect that pupil diameter is also modulated by individuals' volatility estimates. In the tasks implemented by Preuschoff et al. and Satterthwaite et al., a single probabilistic context was used, meaning there was no inherent volatility. As such, it was not possible to investigate any impact of volatility on pupil diameter. In de Berker et al.'s task, the probabilistic relationship between cues and outcomes was unstable. By applying the HGF model to individuals' behavioural data, it was possible to capture their trial-wise volatility estimates and to isolate them from their estimates of surprise and uncertainty. However, any impact of volatility on pupil diameter was not the focus of the study and was therefore unaddressed.

Nevertheless, pupillometric measures taken under alternative task paradigms have suggested that changes in pupil diameter are linked to estimates of volatility. In an experiment by Nassar et al., participants undertook an isoluminant predictive-inference task in which they were required to make trial-wise predictions about the next number in a series (Nassar et al., 2012). Numbers were drawn from a Gaussian distribution. The mean of the Gaussian distribution was stable within a particular context but changed unpredictably over time, introducing volatility. The standard deviation of the distribution introduced noise, i.e., random fluctuations in the data generated by an otherwise stable context. Over each block of trials, the standard deviation of the distribution was set to either 5 or 10, introducing low or high noise, respectively.

Transient increases in pupil diameter (which the authors describe as pupil changes) were augmented under high surprise (i.e., high sensory prediction error) but low noise. Given

that larger prediction errors would be expected following a change in context, it was suggested that pupillary dilation might reflect perceived contextual instabilities arising at change-points (Figure 1.6A). Accordingly, transient increases in pupil diameter were found to predict change-point probability (Figure 1.6B), and coincided with an increase in learning rate that would be expected to accompany a contextual change. Together, these findings indicate that pupil diameter is sensitive to sensory changes that arise due to environmental volatility.

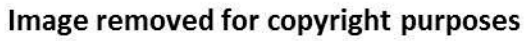

Figure 1.6 Relationship between post-outcome pupil change, prediction error, noise and change-point probability. *(A) Transient pupil change increased as a function of prediction error magnitude, scaled as a function of noise (black = low noise; grey = high noise). Pupil change was computed as the difference in mean (z-scored) pupil diameter measured late (time = 1-2 seconds) versus early (time = 0-1 seconds) post-outcome. (B) Increased pupillary dilation predicted a higher change-point probability, as estimated by a Bayesian learning model. Data are mean ± SEM. Figure adapted from Nassar et al., 2012.*

In addition, average pupil diameter during the outcome viewing period reflected the reliability with which recent trial history indicated the current contextual relationship (Figure 1.7), a parameter that the authors frame as relative uncertainty. Relative uncertainty encompasses my working definitions of irreducible and estimation uncertainty, reflecting both the unreliability with which a single sample can be predicted from a distribution with a known mean and the unreliability of an individual's current estimate of that mean. It increases rapidly after a change-point and then decreases as more data are observed from the current distribution and an individual learns the rules

of the current context. The finding that average pupil diameter mirrored relative uncertainty, peaking after a change-point and then decreasing over subsequent trials (Figure 1.4), echoes, at least in part, our previous finding of increased pupil diameter under increased irreducible uncertainty (de Berker et al., 2016).

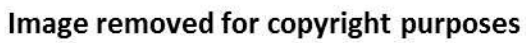

Figure 1.7 Relationship between average pupil diameter and relative uncertainty. *(A) Average pupil diameter as a function of trials pre- and post-change points. Pupil average is the mean (z-scored) pupil diameter across a 2 second outcome viewing period. The asterisk indicates that average pupil diameter on the trial immediately following the change-point was significantly greater than on all other trials. (B) An increase in average pupil diameter predicted an increase in relative uncertainty, as estimated by a Bayesian learning model. There was no relationship between relative uncertainty and transient pupil change. Data are mean ± SEM. Figure adapted from Nassar et al., 2012.*

Importantly, adaptive behaviour in dynamic probabilistic environments requires that the relative impact of incoming sensory information is modulated in line with different sources of uncertainty. When irreducible and estimation uncertainty (i.e., relative uncertainty) arise from noise, the average perceptual estimate over all historical sensory data is most predictive of future observations. In contrast, when volatility uncertainty arises from a change in probabilistic context, only the most recent observations are relevant. Thus, historical sensory data should be discounted and beliefs should be updated rapidly, in accordance with incoming sensory data, to maximise prediction accuracy. Nassar et al.'s finding that pupil diameter can predict both change-point probability and relative uncertainty suggests that the pupil might reflect the dynamics of the hypothesised noradrenergic processes that underlie learning in uncertain and volatile environments.

Recently, Browning et al. conduced an explicit investigation into the relationship between pupil diameter and volatility (Browning et al., 2015). In an aversive probabilistic learning

task, participants were required to select one of two visual cues, each of which was probabilistically linked to a subsequent electrical shock of a magnitude defined by that cue. The probability that each cue predicted an electrical shock changed over time, introducing volatility. During a stable block, the cue:outcome probabilities were fixed for 90 trials. During a volatile block, the cue:outcome probabilities switched every 20 trials. Pupillary dilation following outcome presentation was modulated by both surprise (Figure 1.8A) and volatility (Figure 1.8B), as estimated by a Bayesian learning model.

Figure 1.8 The effects of surprise and volatility on post-outcome pupil diameter during aversive probabilistic learning. *(A) Pupil dilation was modulated by surprise 1-3 seconds post-outcome and (B) by volatility 2-5 seconds post-outcome. Data are mean beta weights ± SEM. Figure adapted from Browning et al., 2015.*

Moreover, individuals with high trait anxiety demonstrated a reduced ability to adjust their learning rate following switches from stable to volatile blocks. Specifically, it seems that these individuals could not increase their learning rate under volatility, instead showing equivalent learning in both stable and volatile blocks. Elevated trait anxiety was also associated with a decreased mean pupillary response to volatility. There was no modulatory effect of trait anxiety on the mean pupil response to surprise. Given that anxiety has been linked to dysfunctional noradrenergic neurotransmission (Gorman et al., 2001), these results are compatible with the theorised functional relationship between noradrenergic neurotransmission, pupillary dynamics and learning under environmental volatility.

In sum, varied behavioural and pupillometric evidence suggests that pupil diameter is modulated by an individual's perceptual estimates under environmental uncertainty, with pupil dilation having been linked to uncertainty, surprise and volatility. Given the proposed role for NA in learning under volatility uncertainty, and the electrophysiological, pharmacological and neuroimaging data that suggest that pupil diameter is sensitive to

NA neurotransmission, the notion that pupil dilation can be used as a proxy for dynamic noradrenergic uncertainty computations is appealing. However, the foregoing investigations of pupillary responses to perceptual estimates have been heterogeneous: they have used different behavioural paradigms that exposed participants to different forms of uncertainty, and probed the impact of different combinations of perceptual beliefs on pupil diameter. As such, it is difficult to isolate the contribution of particular perceptual estimates to pupil diameter with confidence. Therefore, in **Chapter 6**, I combine a probabilistic learning task, pupillometry and a hierarchical Bayesian learning model to assess the impact of irreducible uncertainty, surprise and volatility on pupil diameter. Further, by utilising two pharmacological manipulations of NA, I causally assess whether any pupillary responses to these perceptual beliefs are under dynamic noradrenergic modulation.

1.6 Thesis overview

To summarise, this thesis addresses the neuromodulatory underpinnings of learning and action in uncertain environments. In a series of experiments, I seek to identify the relative contributions of NA, ACh and DA to perceptual belief updating and motor response modulation within a unified computational framework of irreducible, estimation and volatility uncertainty. Specifically, I hypothesised that:

1. Pharmacologically manipulating NA would modulate learning under volatility uncertainty arising from environmental instability;
2. Pharmacologically manipulating ACh would modulate learning under estimation uncertainty arising within probabilistic environmental contexts;
3. Pharmacologically manipulating DA would modulate the sensitivity of motor responses to perceptual estimates of sensory prediction error;
4. Inter-individual variations in *COMT* genotype would alter DA neurotransmission and thus the sensitivity of motor responses to perceptual estimates, echoing the effect of pharmacologically manipulating DA;
5. Pharmacologically manipulating NA would modulate dynamic computations of uncertainty arising from environmental volatility, as reflected by pupil diameter.

In **Chapter 2**, I introduce the methodological techniques implemented to interrogate the relative contributions of NA, ACh and DA to perceptual belief updating and response modulation under different forms of uncertainty.

In **Chapter 3**, I focus on the HGF model I apply to behavioural data in the following experimental chapters. I introduce the HGF as a generic Bayesian framework of individual learning under irreducible, estimation and volatility uncertainty. Further, I describe how I extended the original instantiation of the HGF so that it was possible to capture both individual learning and response modulation under these three forms of uncertainty.

The empirical work of this thesis is presented in **Chapters 4 to 6:**

In **Chapter 4**, I implement pharmacological interventions in 128 healthy human participants to characterise the influences of noradrenergic, cholinergic and dopaminergic receptor antagonism on individual computations of uncertainty during a probabilistic serial RT task. Using the novel instantiation of the HGF, I disentangle the effects of the neuromodulators on individual perceptual belief updating from any effects on the sensitivity of motor responses to perceptual estimates.

In **Chapter 5**, I adopt a behavioural genetics approach to probe deeper into the role of DA in learning and response modulation under uncertainty. Specifically, I use the same serial probabilistic RT task and the same HGF model to assess the impact of *COMT* genotype on learning and action in 116 healthy human participants. As such, I aim to determine whether natural inter-individual variations in cortical dopaminergic neurotransmission modulate individual perceptual belief updating and/or the sensitivity of motor responses to perceptual estimates.

In **Chapter 6**, I focus on the role of NA in learning under uncertainty. Combining pupillometry, the HGF model and two pharmacological manipulations of NA in 90 healthy human participants, I characterise dynamic noradrenergic responses to uncertainty, surprise and volatility during an auditory probabilistic learning task.

Finally, in **Chapter 7**, I discuss the implications of this work, drawing together insights from the different lines of research presented in this thesis.

1.7 Acknowledgement of contributions

I gratefully acknowledge Christoph Mathys' assistance in developing the novel instantiation of the HGF presented in **Chapter 3**. I thank Joseph Galea for coordinating the behavioural testing sessions for **Chapter 5**. I also thank Graziella Quattrocchi, Simon Little and Diane Ruge for providing clinical supervision during the pharmacological experiments presented in **Chapters 4** and **6**. Genotyping for **Chapter 5** was performed by the West Midlands Genetic Laboratory at the Birmingham Women's Hospital.

2 Methods

In this chapter, I introduce the methodological techniques implemented in this thesis to investigate the neuromodulatory bases of learning and action in uncertain environments. I describe how the methods were developed and summarise the experimental considerations that were made.

2.1 Behavioural paradigms to investigate learning and action under uncertainty

In order to interrogate the roles of NA, ACh and DA in computing *irreducible uncertainty*, *estimation uncertainty* and *volatility uncertainty*, and in modulating motor responses to uncertainty, it was necessary to design behavioural paradigms that exposed participants to each quantity. For **Chapters 4** and **5**, I developed a novel probabilistic serial reaction time task that required participants to track these three forms of uncertainty to engender fast, accurate responses to visual stimuli. For **Chapter 6**, I adapted a previously documented probabilistic learning task (den Ouden et al., 2010; Iglesias et al., 2013; de Berker et al., 2016) that required participants to make accurate predictions about trial outcomes given uncertain cues whose reliability changed over time.

2.1.1 Probabilistic serial reaction time task

2.1.1.1 Task design

The probabilistic serial reaction time task (PSRTT) required participants to respond to the trial-wise presentation of one of four visual stimuli by pressing an appropriate button as quickly as possible. At any given time, the trial sequence was generated by one of eight transition matrices (TMs), which changed every 50 trials. In each case, there were 16 combinations that determined the probabilistic relationship between the stimuli presented on the current trial *t*, and the previous trial, *t-1*. In **Chapter 4**, three types of TM were utilised: 1st-order sequences, Alternating sequences, and 0th-order sequences (Figure 2.1).

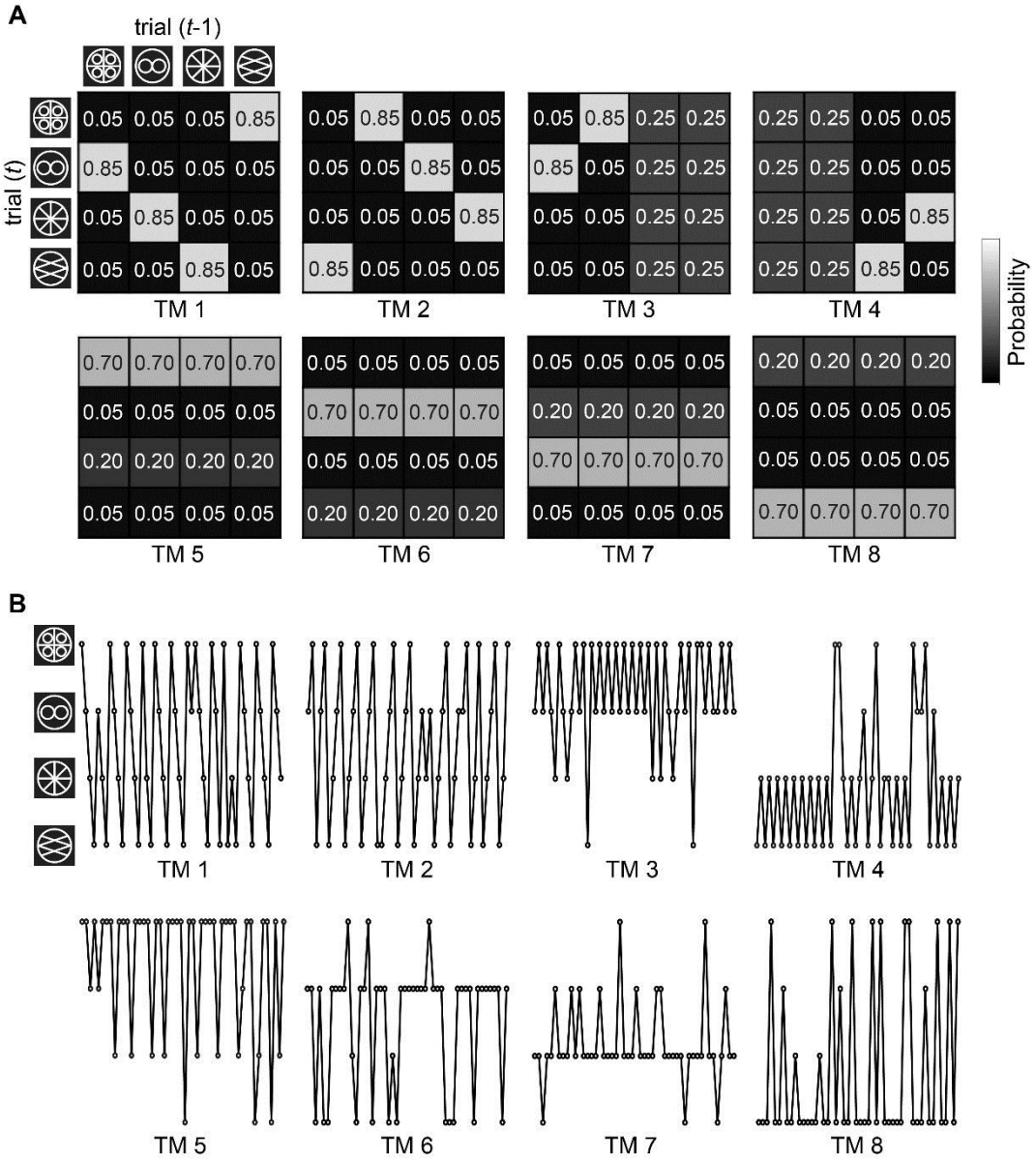

Figure 2.1 Probabilistic structure of the PSRTT. *(A) The eight TMs that determined the probabilistic relationship between the visual stimulus presented on the current trial, t, and that presented on the previous trial, t-1. (B) Example trial sequences generated from the eight TMs. TMs 1 and 2 generated 1st-order stimulus sequences in which there was a high probability of the sequences 1-2-3-4 and 4-3-2-1 occurring respectively. TMs 3 and 4 resulted in a high probability of alternating between two stimuli. TMs 5-8 were 0th-order sequences that led to one stimulus occurring with a high probability, one with a mid probability and two with a low probability. The TM switched to a different TM every 50 trials. Over the course of the experiment, each TM occurred multiple times in a pseudorandom order, with no consecutive repeats of any one TM. The overall probability of each stimulus was equal across 1200 trials.*

In **Chapter 5**, the task design was simplified slightly in that the Alternating TMs were replaced with two additional 1st-order sequences (Figure 2.2). All other TMs were identical to those used in **Chapter 4**.

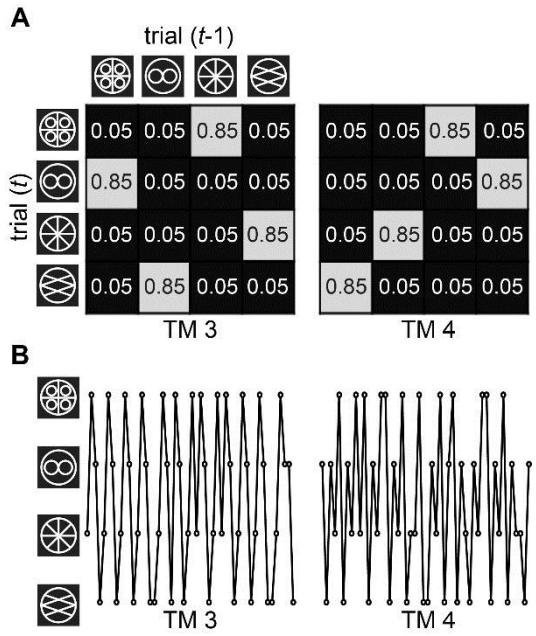

Figure 2.2 Alternative probabilistic contexts. *(A) In* ***Chapter 5****, two additional 1st-order TMs replaced the Alternating TMs used in* ***Chapter 4****. (B) Example trial sequences indicate that TM 3 resulted in a high probability of the stimulus sequence 1-2-4-3 occurring, and TM 4 a high probability of the stimulus sequence 1-4-2-3. In* ***Chapter 5****, TMs 1, 2, and 5-8 were identical to those defined in Figure 2.1.*

In both versions of the task, the order of the TMs was pseudorandom, with no consecutive repeats. This pseudorandom order of TMs was used to generate one stimulus sequence that was used for all participants to ensure comparable learning processes and model parameter estimates between individuals. Importantly, the overall probability of each stimulus was equal across the complete set of trials.

The task design created transient contexts that participants could infer from stimulus observations, allowing them to reduce their uncertainty about events before they occurred (Harrison et al., 2006). Nonetheless, the probabilistic nature of these contexts also produced unexpected stimulus outcomes, i.e., a sensory prediction error (PE). For fast and accurate responses, participants had to track *irreducible uncertainty* arising from the inherent randomness of the probabilistic transitions between consecutive stimuli; *estimation uncertainty* arising from their imperfect knowledge of the probabilistic relationships governing stimulus transition contingencies within contexts; and *volatility uncertainty* maintained by the unsignalled contextual instability.

By implementing a hierarchical Bayesian learning model, with a response model adapted for this experimental paradigm (**Chapter 3**), it was possible to map an individual's beliefs about stimulus transitions, transition contingencies, and volatility, and the respective *irreducible*, *estimation* and *volatility* uncertainty about these beliefs, onto his/her observed reaction time (RT) responses. In **Chapter 4**, the behavioural paradigm and computational modelling were combined with pharmacological manipulations of NA, ACh and DA in order to characterise the roles of the three neuromodulators in perceptual belief updating under these three forms of uncertainty, and in modulating motor responses to this uncertainty. In **Chapter 5**, genetic analyses of the Val^{158}Met polymorphism in the *COMT* gene were used to probe the impact of inter-individual differences in dopaminergic neurotransmission on learning and action under uncertainty, thereby extending the findings of the pharmacological study.

2.1.1.2 Training

Each stimulus was associated with one particular button. Each participant acquired the stimulus-response mappings during a training block in which they received visual error feedback after each trial. The training session comprised at least 100 trials and did not finish until participants had reached a minimum performance criterion of 85% accuracy on the last 20 trials. This was to ensure that any learning during the task was related to the probabilistic relationships governing stimulus transitions, rather than the stimulus-response mappings themselves.

2.1.1.3 Button boxes

Participants made their speeded responses via a custom-made button box. Four button boxes were used so that four participants could be tested in a multiple participant setup (see section 2.1.3). Two of the button boxes transmitted data to the testing computer via the serial port, and two via the parallel port. Button boxes, rather than computer keyboards, were used for improved precision of the logged RTs. Both the serial and parallel port button boxes had a temporal precision of 3-13ms, compared to a keyboard's 33ms.

2.1.2 Probabilistic learning task

2.1.2.1 Task design

The probabilistic learning task (PLT) was closely modelled on a paradigm used in three recent studies (den Ouden et al., 2010; Iglesias et al., 2013; de Berker et al., 2016). On each trial, participants were presented with one of two auditory cues: a low pitch (450Hz)

or high-pitch (1000Hz) tone. They were required to make a prediction, signalled with a speeded button press, as to which auditory outcome (the word “cow” or the word “pig”) was likely to follow. A probabilistic mapping between stimulus and outcome exposed participants to *estimation uncertainty* about the current cue:outcome relationship. This probabilistic mapping shifted over the course of the experiment, introducing *volatility uncertainty* and requiring participants to constantly track the cue:outcome relationship over time. The task’s probabilistic nature also meant that unlikely outcomes were possible on any trial, giving rise to *irreducible uncertainty*. To make accurate predictions, participants therefore had to track three forms of uncertainty throughout.

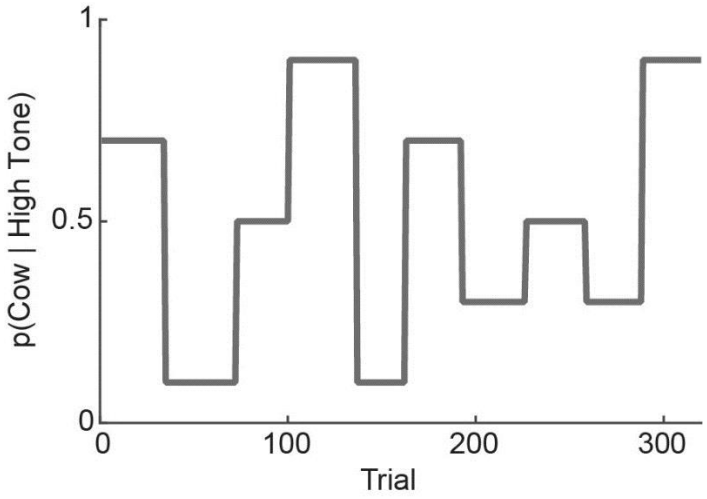

Figure 2.3 The instability of the PLT over time. *The probabilities governing the cue:outcome relationships shifted unpredictably over time, producing fluctuations in uncertainty. The probabilities governing each block varied from heavily biased (0.9/0.1 and 0.1/0.9), through moderately biased (0.7/0.3 and 0.3/0.7) to unbiased (0.5/0.5).*

2.1.2.2 Auditory stimuli

Since the PLT was combined with pupillometry, the visual stimuli used in previous iterations of this task (den Ouden et al., 2010; Iglesias et al., 2013; de Berker et al., 2016) were replaced with auditory stimuli so as to eliminate any effects of luminance changes on pupil diameter. Moreover, to avoid any inter-stimulus difference in auditory saliency effects on the pupil, participants underwent an adaptive, two-alternative forced choice procedure before undertaking the behavioural task so as to match the subjective loudness of the auditory cues and outcomes. The two auditory cues and two auditory outcomes had the same durations (300ms and 600ms respectively). The auditory outcomes were neutral words that belonged to a single (animal) category and were easy to distinguish.

2.1.2.3 Training

Participants were trained on the PLT before starting it. During four training blocks of five trials each, participants familiarised themselves with making predictions by button press on presentation of auditory cues. Participants were told at the start of each training block which cue and which outcome would be presented on each of the following five trials. After the outcome was presented, they were provided with visual error feedback. Each combination of cue and outcome was presented across the four training blocks. The order of the four training blocks (i.e., the pairings between each cue and each outcome) was counterbalanced across participants. To familiarise themselves with the timings of the PLT, participants then completed 12 practice trials without error feedback. On each trial there was a 50% probability that either cue would be followed by either outcome.

2.1.3 Multiple participant setup

For efficiency, I implemented multiple participant testing (Figure 2.4A). For the PSRTT, four participants were tested simultaneously. To ensure that they were not distracted while completing experimental tasks, participants wore ear defenders and sat in individual booths. For the PLT, which was combined with pupillometry, experimental sessions were staggered such that a second participant arrived once the first participant had received their drug or placebo and had entered the 1.5 hour waiting period that preceded commencement of the main behavioural task (see section 2.3.4 for details). To ensure that all participants received identical task instructions, they were provided in written form.

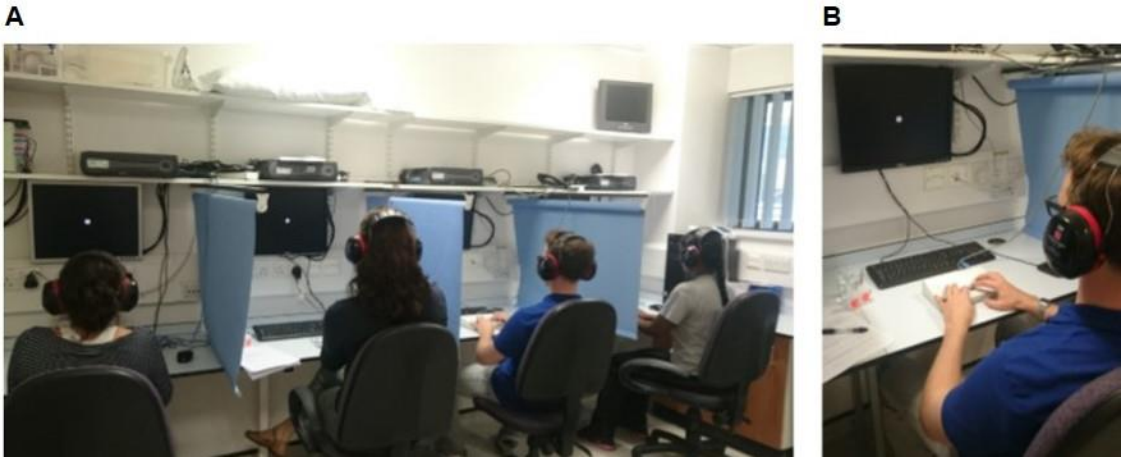

Figure 2.4 Behavioural setup used during the PSRTT. *(A) A multiple participant setup was employed for simultaneous testing of four participants. (B) Each participant sat fixating a central white cross presented on a black computer screen positioned 60cm away. They were instructed to place their left and right index and middle fingers on the*

four buttons of a custom-made button box, and to maintain this position throughout the task. Following the trial-wise presentation of one of four visual stimuli, participants were required to make a speeded-button press response.

2.2 Computational modelling of learning and action

Computational modelling makes it possible to estimate the inferences made by participants during a behavioural task. Hierarchical Bayesian models have proven powerful for explaining the adaptation of behaviour to probabilistic contexts in dynamic environments (Behrens et al., 2007; den Ouden et al., 2010; Nassar et al., 2010; Wilson et al., 2013). In particular, the Hierarchical Gaussian Filter (HGF) model (Mathys et al., 2011, 2014) has been successfully applied in several recent studies of probabilistic learning under volatility (Iglesias et al., 2013; Diaconescu et al., 2014, 2017; Hauser et al., 2014; Vossel et al., 2014a, 2014b, 2015; de Berker et al., 2016).

The HGF models an individual's learning across three levels: it tracks beliefs about environmental events (e.g. the presentation of a stimulus) at level 1, the probabilistic relationships linking different environmental events at level 2, and the volatility of these relationships at level 3. As such, it is possible to access the respective *irreducible*, *estimation* and *volatility uncertainty* about these beliefs, and to infer an individual's beliefs about the causes of his/her sensory inputs.

The HGF is hierarchical in that learning not only occurs simultaneously at multiple levels, but that belief updating at one level is constrained by beliefs at the level above. This provides a generic framework for implementing learning rates, which are crucial for learning in volatile environments (Behrens et al., 2007; den Ouden et al., 2010). Importantly, the HGF does not assume fixed "ideal" learning across individuals but rather contains participant-specific parameters that couple the hierarchical levels and allow for individual expression of (approximate) Bayes-optimal learning.

The PLT used in **Chapter 6** was compatible with a pre-existing instantiation of the HGF (Iglesias et al., 2013; de Berker et al., 2016). The PSRTT used in **Chapters 4** and **5** required the development of a novel instantiation of the HGF. This new version has two components: a three-level *perceptual model* of an agent's mapping from environmental causes to sensory inputs, and a *response model* that maps those inferred environmental causes to observed RT responses. Full details are presented in **Chapter 3**.

2.3 Pharmacological manipulations of neuromodulatory function

In **Chapters 4** and **6**, I employed pharmacological interventions that manipulate the noradrenergic, cholinergic and dopaminergic neuromodulatory systems in order to causally asses the roles of NA, ACh and DA in learning and response modulation under uncertainty.

In **Chapter 4**, I utilised three antagonists for the different neuromodulators, and compared learning and action under each manipulation to placebo. Prazosin acts as a (nor)adrenergic antagonist. Specifically, the drug is an inverse agonist that binds to α1-adrenoceptor but induces a pharmacological effect opposite to that of the receptor's endogenous ligand, (nor)adrenaline (Zhu et al., 2000). Biperiden acts as an antagonist at cholinergic muscarinic M1-receptors. Since the drug has affinity, but no efficacy, for M1-receptors, it competes with ACh for receptor binding sites and dampens the effect of the natural cholinergic ligand. Similarly, when administered at a sufficient dose, haloperidol acts as a dopaminergic antagonist, competing with DA for D1- and D2-receptor binding sites and dampening the pharmacological effect of DA.

In **Chapter 6**, I probed dynamic noradrenergic responses to uncertainty by using two pharmacological agents that have bidirectional effects on the NA system. As in **Chapter 4**, I used prazosin to antagonise NA. In addition, the selective NA reuptake inhibitor, reboxetine, was used to upregulate the NA system by blocking the action of the noradrenaline transporter (NET), which is responsible for the reuptake of extracellular NA.

2.3.1 Between-subjects design

For each pharmacological experiment (**Chapters 4** and **6**), I implemented a between-subjects design. As such, each participant attended one experimental session during which they received either an active drug or a placebo. The principal reason for this was that my behavioural paradigms involved learning. Since each participant undertook the behavioural task only once, it was possible to eliminate any learning effects that may otherwise have been carried over from previous sessions. Moreover, a between-subjects design made it possible to use a single, pseudorandomly generated trial sequence for all participants undertaking the PSRTT (**Chapters 4** and **5**), ensuring comparable learning processes and model parameter estimates. A further benefit of a between-subjects design was that participants were not required to attend multiple experimental sessions, minimising any problems caused by drop-outs.

2.3.2 Safety

For safety purposes, it was necessary to screen participants to rule out intolerances or contraindications for the active drugs. To reduce the risk of side effects, I aimed to select relatively low drug doses that were nevertheless in line with previous studies showing clear behavioural and neurophysiological effects (Dostert et al., 1997; Ziemann et al., 1997; Meintzschel and Ziemann, 2006; de Martino et al., 2007; Jepma et al., 2010; Korchounov and Ziemann, 2011; Bestmann et al., 2014). In addition, the following exclusion criteria applied: history of neurological or psychiatric disease, intake of medication (other than contraceptives), smoking, regular drug use, baseline blood pressure below 100/60, and current participation in other pharmacological studies. As an additional precaution, heart rate and blood pressure measurements were taken at three timepoints during the experimental session (see 2.3.3). A clinician was also available during each testing session in case of any medical queries or concerns.

2.3.3 Physiological, psychometric and subjective control measures

A between-subjects design meant that it was necessary to match the different drug-groups for physiological, psychometric and subjective variables that could influence drug responses and behaviour during the experimental tasks. At recruitment, the study clinician pseudorandomly assigned participants to receive one of the active drugs or a placebo in order to ensure a balanced distribution of gender, age and body weight. Importantly, a double-blind design was achieved since both the experimenter (L.M.) and the participants were blind to the drug conditions. Since nicotine acts at cholinergic (nicotinic) receptors, smokers were excluded.

Since participants' working memory, impulsivity, risk-taking and distractibility could influence behaviour during the tasks, participants undertook computerised versions of the Digit Span test, Barratt Impulsiveness Scale (BIS-11) (Patton et al., 1995), Domain-Specific Risk-Taking (DOSPERT) Scale (Blais and Weber, 2006) and Cognitive Failures Questionnaire (CFQ) (Broadbent et al., 1982) at baseline (i.e., before taking a drug). In both pharmacology experiments (**Chapters 4** and **6**), scores on these tests were established to be equivalent across drug-groups.

2.3.3.1 Digit span

Digit span is a common measure of working memory. Participants were instructed to memorise sets of digits that were presented to them via stereo headphones, and then to repeat those digits, in the order in which they had been read, by typing them via a computer keyboard. The first trial started with a set of three digits. Every two trials, the

set-length increased by one digit. For each set-length, participants were required to repeat at least one of the two sets correctly to progress to the next level. The task finished either when a participant incorrectly repeated two sets of the same length, or when they correctly repeated the maximum set-length of nine digits. After completing the forward digit span, participants undertook the backwards digit span. Here the rules were the same, except that participants had to repeat the digits that were presented to them in the reverse order.

2.3.3.2 Barratt impulsiveness scale (BIS-11)

The BIS-11 (Patton et al., 1995) consists of 30 items describing common impulsive, or non-impulsive (for reverse scored items), behaviours and preferences. It interrogates attentional, motor and non-planning impulsiveness. Participants used a 4-point scale to self-report whether they engaged in particular behaviours, or had particular preferences, “rarely/never”, “occasionally”, “often”, or “almost always/never”. The higher the summed score for all items, the higher the level of impulsiveness.

2.3.3.3 Domain-specific risk-taking scale (DOSPERT)

The DOSPERT is a validated scale (Blais and Weber, 2006) that assess an individual’s tendency for risk-taking behaviours, and their perceived-risk attitudes (defined as the willingness to engage in a risky activity as a function of its perceived riskiness) in five domains: ethical, financial, health/safety, social, and recreational decisions. It consists of 30 items. Participants self-reported the likelihood that they would engage in a described activity or behaviour, or how risky they considered a described behaviour to be, using a 7-point scale ranging from “extremely unlikely” to “extremely likely”, and from “not at all risky” to “extremely risky”, respectively. A higher total score indicates higher risk-taking behaviour.

2.3.3.4 Cognitive failures questionnaire (CFQ)

The CFQ (Broadbent et al., 1982) measures self-reported failures in perception, memory and motor function, and thus can be used to approximate an individual’s distractibility. It consists of 25 questions, such as “Do you read something and find you haven’t been thinking about it and must read it again?” Participants responded using a 5-point scale: “very often”, “quite often”, “occasionally”, “very rarely” and “never”. A higher score indicates a higher failure rate.

2.3.3.5 Controlling for subjective measures

To assess any subjective drug effects, participants used sixteen visual analogue scales (VAS) to self-report their mood at baseline, before starting the main behavioural task (i.e., when the active drugs were at their most active) and after completing the behavioural task. For each VAS, two extreme moods, such as *alert/drowsy* and *tense/relaxed* were presented at either end of a 100-point scale and participants had to mark their current subjective feeling. The 16 measures were used to calculate scores for alertness, calmness and contentedness (Bond and Lader, 1974). Since there was a significant interaction between alertness and drug in both pharmacology studies, an alertness covariate was included in the analyses (see **Chapters 4** and **6** for details).

For completeness, heart rate and blood pressure measurements were taken at the same three timepoints as the VAS scores so as to monitor any physiological drug effects.

2.3.4 Drug administration times

Drug administration times were selected such that participants undertook the main experimental task when the active drug they had been administered was at its most active. Since average time-to-peak plasma concentrations varied across drugs, two different drug administration times were used. In the first pharmacology study (**Chapter 4**), haloperidol was administered two hours before the main experimental session, while prazosin and biperiden were administered 1.5 hours in advance (Ziemann et al., 1997; Meintzschel and Ziemann, 2006; Korchounov and Ziemann, 2011). A random 50% of participants from the Placebo group were administered a placebo tablet at the first timepoint, and the other 50% at the second timepoint.

In the second pharmacology experiment (**Chapter 6**), a minor modification was made to this approach so that all participants underwent an identical administration schedule. This time all participants were administered two tablets thirty minutes apart. Reboxetine was administered two hours before the main experimental task and prazosin 1.5 hours in advance (Dostert et al., 1997; de Martino et al., 2007; Jepma et al., 2010). At the timepoint at which participants in the Reboxetine and Prazosin groups did not receive an active drug, they were administered a placebo tablet. Participants in the Placebo group were administered a placebo tablet at both timepoints. This method ensured that, in multiple participant testing sessions, individuals would not infer any differences in how the different participants were being treated, and ensured that the experimenter (L.M.) remained blind to the drug conditions.

In both experiments, the study clinician administered the drug or placebo while the experimenter was away from the testing room. For comparable metabolism rates, participants were asked not to eat for at least one hour before the first drug administration time.

2.4 Behavioural genetics

An alternative approach to studying neuromodulatory systems is to interrogate genetic polymorphisms that introduce natural inter-individual variations in neuromodulatory function. A polymorphism of particular interest in the behavioural genetics literature is the $Val^{158}Met$ polymorphism in the *COMT* gene, which modulates activity of the COMT enzyme and thus DA neurotransmission. In **Chapter 5**, I combined the same PSRTT and novel instantiation of the HGF applied in **Chapter 4** to assess individual perceptual belief updating and response modulation as a function of *COMT* genotype. As such, it was possible to probe whether inter-individual variations in DA neurotransmission were associated with altered learning and/or action under uncertainty.

2.5 Pupillometry

As discussed in the Introduction, the neuroscience literature has previously linked subjective uncertainty computations to changes in pupil diameter (Preuschoff et al., 2011; Nassar et al., 2012; de Gee et al., 2014; de Berker et al., 2016), and pupil dilation to noradrenergic neuronal activity in the locus coeruleus (Aston-Jones and Cohen, 2005a; Varazzani et al., 2015; Joshi et al., 2016). In **Chapter 6**, I combined pharmacological manipulations of NA with a behavioural paradigm and pupillometry to assess the neuromodulator's causal impact on learning in dynamic environments and on pupillary responses to uncertainty.

While participants undertook the PLT, the diameter of the left pupil was measured using an infrared ASL Eye-Trac 6 System (Applied Science Laboratories, USA), sampling at 120Hz (Figure 2.5). In order to minimise movement, participants sat with their head supported in a forehead- and chin-rest. The viewing distance was fixed at 60cm. Participants were instructed to maintain fixation on a horizontally-centred fixation cross presented on an isoluminant, grey screen. The vertical position of the fixation cross was such that the participant's line of vision was straight ahead. Auditory stimuli were presented via stereo headphones.

The eyetracker system calculates pupillary gaze by measuring the distance between the location of a participant's pupil and corneal reflection (CR). For each participant, the eyetracker was calibrated to account for inter-participant differences in the relationship between the pupil and CR. The central calibration point was positioned at the location of the centre-point of the fixation cross used during the behavioural task. Calibration was repeated after each rest period to adjust for any subtle differences in head position. In order to align the pupil diameter time-course with experimental events occurring in the behavioural task (i.e., the precise timing of cue, response and outcome onsets) triggers were sent via the testing computer's parallel port to the eyetracker system.

Figure 2.5 Pupillometry setup used during the PLT. *Participants sat in a darkened room with their head supported by a forehead- and chin-rest. During the experimental task, participants were instructed to maintain fixation on a black cross horizontally centred on a grey isoluminant display positioned 60cm ahead. The fixation cross was presented parallel with the participant's line of vision. An infrared eyetracker was used to measure the diameter of the left pupil. Auditory stimuli were presented via stereo headphones. Participants were instructed to position their left and right index fingers on two marked keys on a computer keyboard, and to maintain this position throughout the experiment. On the presentation of each trial-wise auditory cue, participants were required to make a prediction via a speeded button-press as to which auditory outcome they thought would follow.*

3 Modelling individual learning and action under uncertainty

*This chapter is based on work presented in **Marshall L**, Mathys C, Ruge D, de Berker AO, Dayan P, Stephan KE & Bestmann S. (2016) Pharmacological fingerprints of contextual uncertainty. PLOS Biology. 14(11): e1002575.*

3.1 Abstract

A mechanistic understanding of perceptual belief updating and response modulation requires specification of the computational principles by which learning and action occur, and identification of their neurophysiological implementation in the brain. As such, the development of computational models is key to elucidating the neurophysiological bases of learning and action in uncertain environments. In 2011, Mathys et al. introduced the Hierarchical Gaussian Filter (HGF) model as a generic hierarchical Bayesian framework for individual learning under various forms of uncertainty inherent in the environment. In this chapter, I first provide a summary of the computational principles that inspired the HGF's construction. Second, I describe how the original instantiation of the HGF is designed to capture individual learning under uncertainty. Third, I focus on a novel extension of the HGF model designed to link perceptual beliefs to action execution. Specifically, this new instantiation of the HGF characterises individual learning and response modulation during the serial probabilistic reaction time task (PSRTT) applied in Chapters 4 and 5.

3.2 Introduction

Computational models offer a sophisticated means by which to probe the brain's mechanisms of learning and adaptive behaviour in complex, uncertain environments. Such models commonly conceptualise the environment as a set of unobservable, hidden states whose temporal dynamics generate our observable, sensory input. Learning about the hidden states that define our present environmental context, and adapting to contextual changes, requires us to update our beliefs about the world by integrating our top-down, experience-driven expectations and our current bottom-up sensory evidence (Yu and Dayan, 2003). In so doing, we can exploit our past experience of the world while also taking into account the current state of the environment, thus improving our ability to predict future events and to prepare adaptive motor responses. One strategy for investigating human learning and action under uncertainty is to formulate the computational principles that underlie perceptual belief updating and response modulation, and identify the neurophysiological implementation of those computations in the brain (Daunizeau et al., 2010a). Indeed, combining formal models of learning and behaviour with the measurement of neural signals has brought significant advances to behavioural neuroscience in recent years (Daw et al., 2011; Takahashi et al., 2011; Iglesias et al., 2013).

3.2.1 Capturing the computational principles that underlie learning and action

Two classes of models that have been applied in an effort to infer the computational underpinnings of learning and action within the brain take inspiration from reinforcement learning and Bayesian learning principles.

3.2.1.1 Reinforcement learning models

Reinforcement learning (RL) models seek to capture how individuals learn to optimise their behaviour in a given environment by predicting the consequences of their actions (Sutton and Barto, 1998; Dayan and Niv, 2008). Inspired by classical conditioning (Peterson, 2004), RL rests upon the notion that humans and other animals learn when an event deviates from its expectations, i.e., it is surprising. As such, under RL models, learning (i.e., belief updating) is driven by prediction errors (PEs), which are formalised as the difference between predicted and actual outcomes. By computing PEs, individuals are proposed to use their experience of the environment to construct an internal model of the associations between stimuli, actions and rewards. Individuals select appropriate actions by searching this internal model space, thereby facilitating the execution of adaptive motor responses to environmental events. Within this framework, an individual

can learn to predict environmental outcomes based on sensory cues, and engender actions that will maximise their chance of gaining reward and minimise their chance of punishment.

A particularly influential RL model has been that constructed by Rescorla and Wagner (Rescorla and Wagner, 1972). It prescribes that beliefs are updated in relation to an individual's pre-existing belief and their current PE, weighted by a learning rate:

$$\mu_n = \mu_{n-1} + \alpha(x_n - \mu_{n-1}),$$

Equation 3.1

where μ_n is the current belief, μ_{n-1} is the belief before making a new observation, α is the learning rate, and $(x_n - \mu_{n-1})$ is the prediction error, i.e. the difference between the new observation, x_n and the existing belief, μ_{n-1}.

The key advantages of the Rescorla-Wagner (RW) model are its conceptual simplicity and computational efficiency; it provides a capable account of how individuals build a primitive model of the world, associating environmental stimuli with their predictors. Moreover, because the learning rate determines the degree to which each PE influences existing beliefs, it modulates the relative influence of recent compared to past events on learning.

However, a fixed learning rate means that individuals will only take into account a fixed number of previous observations when making new beliefs. Since the world is dynamic, and given that changes within different environmental contexts occur at different rates, individuals require a means by which to account for changes in environmental uncertainty. For instance, noisy but otherwise stable environments require low learning rates, which result in stable beliefs, whereas volatile environments necessitate higher learning rates and more flexible beliefs (Behrens et al., 2007; Nassar et al., 2010). There have been various efforts to extend the RW model by developing an adaptive learning rate that allows for flexible belief updating under varying degrees of environmental uncertainty (Kalman, 1960; Sutton, 1992; Nassar et al., 2010; Mathys et al., 2011; Payzan-LeNestour and Bossaerts, 2011; Wilson et al., 2013). Importantly, by allowing learning rates to vary across individuals and environmental contexts, a flexible RL framework can capture action selection that is adaptive to the changes inherent in dynamic environments.

An alternative instantiation of RL is temporal difference (TD) learning. TD models extend the concept that PEs drive learning by incorporating an additional feature of learning commonly included in engineering algorithms, that of prediction (Sutton and Barto,

1981). Specifically, TD models utilise changes, or differences, in predictions over time to drive learning. This means that, whereas RW models describe the association between a predictor and an immediate outcome, TD learning assumes that individuals seek to find the earliest valid predictor of a variable of interest and continually adjust predictions in light of new evidence. A particular focus of TD learning has been capturing how individuals make optimal predictions so that they can select actions that will maximise their cumulative future reward (Sutton and Barto, 1998).

Importantly, there is evidence to suggest that RL may be implemented neuronally. For instance, the signalling of PEs required for RL (Sutton, 1988) has been linked to the phasic activity of DA neurons (Montague et al., 1996). In particular, RL approaches have been central to postulates about electrophysiological and functional neuroimaging measures of brain activity during reward learning, with DA having been proposed to encode reward PE (Schultz et al., 1997; Montague et al., 2004; O'Doherty et al., 2004; Daw and Doya, 2006; D'Ardenne et al., 2008; Hart et al., 2014; Rutledge et al., 2014, 2015). Alternative lines of research have linked DA to the signalling of sensory PE in the absence of reward, and to consequent influences on action selection (Friston et al., 2012; Galea et al., 2012; Bestmann et al., 2014). In both cases, RL has been influential in guiding hypotheses about behavioural and neural dynamics under different experimental learning paradigms.

In sum, the major benefit of RL methods is their capacity to reduce a daunting problem to a series of simple update equations that are both intuitively appealing and computationally feasible. This approach, and its influence on fields ranging from electrophysiology through to cognitive neuroscience and artificial intelligence, has guided our understanding of learning. RL models also offer a useful computational framework for investigations of anticipatory action selection and adaptive behaviour (Killcross and Coutureau, 2003; Matsumoto and Tanaka, 2004; Balleine, 2005; Dolan, 2007; Rushworth and Behrens, 2008), as well as the neuromodulatory contributions to these processes (Yu and Dayan, 2005; Pessiglione et al., 2006; Doya, 2008). Moreover, their non-normative, descriptive nature allows for modelling aberrant modes of learning that occur in disease states such as schizophrenia or depression (Smith et al., 2006; Frank, 2008; Murray et al., 2008; Dayan and Huys, 2009).

Nonetheless, RL has its limitations. First, from a theoretical perspective, it is a heuristic approach that does not follow from the principles of probability theory that would be expected to support optimal learning. Second, at a practical level, RL often performs badly in real-world situations where environmental states and the outcomes of actions

are not known to the individual but must be inferred or learned. Third, RL models do not permit an explicit representation of uncertainty, which appears a shortcoming. Indeed, a strong line of argument from probability theory suggests that learning would be improved if the brain were to represent beliefs as probability distributions, whose variance inherently capture uncertainty, rather than single quantities (O'Reilly et al., 2012).

To illustrate the advantage of capturing uncertainty for optimal belief updating, we can consider a coin toss. In this scenario, we know that the probability of observing heads on each toss is 0.5. If we bet that the outcome of an upcoming coin flip will be heads, at outcome we experience a (reward) PE, which is positive (+0.5) if the outcome is heads or negative (-0.5) if it is tails. Thus, even in situations where we have a pre-existing model of the environment and there is nothing left to learn, we experience PEs.

If we start with the assumption that the coin is unbiased, this implies that after 100 coin flips, we have learned nothing. We are now very sure that the coin is unbiased but, because RL does not explicitly capture uncertainty, it does not offer a means to represent this confidence. This seems wasteful from a neurophysiological perspective; supposing DA signals the PE that follows each coin toss, the neuromodulator will be promoting neuronal plasticity in a situation where this is nothing to left to learn (Yu and Dayan, 2005). Although suggestions have been made as to how RL models might be modified to account for this (Preuschoff and Bossaerts, 2007), it is generally thought that this is a limitation of this class of models (Daw and O'Doherty, 2013). To incorporate an explicit representation of uncertainty, and therefore improve models of learning and behaviour, researchers have turned to Bayesian statistics (Gershman and Niv, 2010).

3.2.1.2 Bayesian learning

The optimal method for learning from new information was first described by Laplace, who set out the laws of inductive inference (Laplace, 1774, 1812). Inductive reasoning is inherently uncertain; it concerns the degree to which a conclusion is credible according to a particular set of evidence. The mechanism by which conditional probabilities are updated in inductive inference was later formally described by Bayes' theorem:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

$$P(A|B) \propto P(B|A)P(A)$$

$$posterior \propto likelihood\ x\ prior$$

Equation 3.2

where *A* and *B* are events and *P(B)≠0*.

As such, inductive inference has become more commonly known as Bayesian inference, or Bayesian learning. Bayesian inference prescribes that the statistically optimal strategy to learning in uncertain environments is to integrate the multiple sources of uncertain information with a weight inversely proportional to their uncertainty. It seems likely that this approach would require an individual's beliefs to be represented as probability distributions (O'Reilly et al., 2012; Pouget et al., 2013). Assuming that these probability distributions are Gaussian, the mean of the distribution would capture an individual's best guess at the value of a particular variable, i.e., its most probable value. The width, or variance, of the distribution would correspond to the uncertainty associated with the representation of that variable. The combination of multiple sources of information produces a probability density function with a variance smaller than that of either of the inputs (Figure 3.1) according to the following equation:

$$\sigma_{total} = \frac{\sigma_1 \sigma_2}{\sigma_1 + \sigma_2},$$

Equation 3.3

where σ_x is the variance associated with the probability distribution x.

Importantly, there is convincing evidence that individuals combine sources of information in this manner. Indeed, in contrast to the aforementioned popularity of RL approaches in studies of reward learning, Bayesian strategies have been more widely applied in the field of sensorimotor control. For instance, researchers have demonstrated that Bayesian integration of this nature is a good predictor of individuals' performance when estimating the height of a bar from noisy visual and haptic information (Ernst and Banks, 2002), when estimating the position of a noise source from visual and auditory cues (Battaglia et al., 2003), and in guiding movements (Körding and Wolpert, 2004).

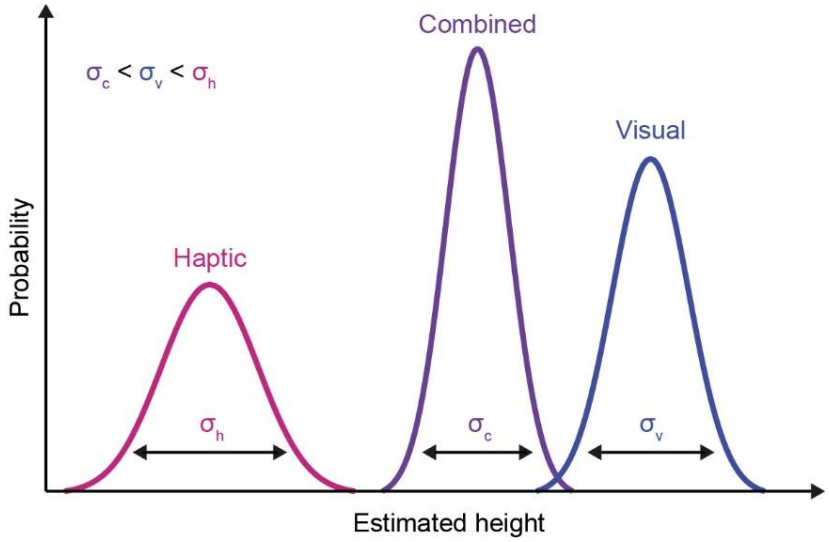

Figure 3.1 Integrating multiple sources of evidence. *Under Bayesian inference, combining multiple sources of uncertain information produces a probability density function with a variance smaller than that of its inputs. For instance, in an experiment conducted by Ernst and Banks, individuals were required to estimate the height of a bar according to noisy visual and haptic information. Integrating these two sources of sensory evidence according to their uncertainty, here represented as the width of a probability density function (σ), produces a combined estimate with a lower variance than either of the two inputs ($\sigma_{combined} < \sigma_{visual} < \sigma_{haptic}$). A model implementing Bayesian integration of this nature was shown to be a good predictor of individuals' performance of the task. Figure adapted from Ernst and Banks, 2002.*

Bayesian inference does not only offer a means for combining multiple sources of uncertain sensory input. As set out in Equation 3.2, Bayes' theorem also provides a way to optimally incorporate current sensory evidence with prior beliefs. To borrow a popular example from Körding and Wolpert, consider a tennis player preparing to return an approaching ball (Figure 3.2). Anticipatory preparation of an adaptive motor response to the ball hitting the ground requires an optimal prediction as to where the ball will bounce. To formulate this prediction, the player must consider both his bottom-up sensory input and his top-down experience-driven beliefs (Körding and Wolpert, 2004). Under ideal Bayesian learning, his *posterior* belief about where the ball is likely to bounce is determined by optimally integrating the *likelihood* of his sensory evidence (i.e., his current visual and auditory input) and his *prior* knowledge of the typical distribution of tennis shots:

$$P(ball\ at\ x|perception\ of\ ball\ at\ x) \propto P(perception\ of\ ball\ at\ x|ball\ at\ x)P(ball\ at\ x)$$

Equation 3.4

The posterior, prior and likelihood are represented as probability distributions, meaning they are inherently associated with an uncertainty. The prior and likelihood are integrated according to their relative uncertainties.

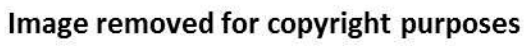

Figure 3.2 Integrating priors and likelihoods over time. *A Bayesian tennis player preparing to return an approaching ball can optimally predict the likely landing location of his opponent's shot by combining the ball's observed trajectory with the prior distribution of shot placement. (A) A schematic of a Kalman filter. At any point in time, the player holds a belief about the state of the world. This belief is updated with a model of the dynamics of the world to calculate the belief at the next point in time. The belief (prior; green patch) is then combined with new sensory information (likelihood; red patch) using Bayes' rule to calculate the belief at the next point in time. (B) To predict the position at which the ball will hit the ground, the player continuously updates his belief in line with incoming sensory information. The posterior at the previous time point is the prior for the current time point. Figure adapted from Körding, 2007.*

Since Bayesian learners integrate their prior knowledge with new information optimally, they would be expected to make better predictions and thus have an evolutionary advantage over non-Bayesian learners. It therefore seems reasonable to hypothesise that the human brain has evolved such that it implements ideal Bayesian learning, i.e., inference on uncertain quantities according to the rules of probability theory (Geisler and Diehl, 2002). At least from a theoretical perspective, this conceptual framework seems well suited to describing information processing in the brain (Knill and Pouget, 2004; Tenenbaum et al., 2006; Körding, 2007). Further, there is considerable evidence from studies on various domains of learning and perception that human behaviour is better

described by Bayesian models than by other theories (Ernst and Banks, 2002; Battaglia et al., 2003; Körding and Wolpert, 2004; Bresciani et al., 2006; Yuille and Kersten, 2006; Behrens et al., 2007; Xu and Tenenbaum, 2007; Orbán et al., 2008; den Ouden et al., 2010), and that human behaviour is close to Bayes optimal in tasks requiring multimodal cue integration and motor adaptation (Yu, 2007; Yu et al., 2009).

Owing to the environment's volatility, the reliability of a prior belief may change over time. In dynamic environments, old beliefs should be rapidly down-weighted relative to new evidence. In stable environments, old information is still valuable. Representing beliefs as probability distributions offers Bayesian learners a means by which to learn and adapt to environmental volatility, with higher uncertainty in dynamic environments reflected by a broader distribution over possible perceptual estimates. In line with this, Behrens et al. demonstrated that humans indeed track environmental volatility, thereby allowing them to learn quickly in dynamic environments, rapidly overwriting old beliefs, but to be more reliant on older information when the environment is stable (Behrens et al., 2007). Learning rates were found to vary systematically with the volatility of an environment's underlying statistical structure, precisely as would be expected in a Bayesian learner. Indeed, comparing the ability of different models to account for individuals' behaviour in volatile environments revealed that an optimal Bayesian learner model outperformed an RL model, despite the latter being tuned to fit the data via free parameters.

However, developing computationally efficient, ecologically valid Bayesian learning models has proven challenging for several reasons. First, in most complex, real-world environmental settings, modelling Bayesian learning involves computationally-demanding, high-dimensional integrals, making online belief updating difficult. An important consideration, therefore, is whether any evolutionary advantage conferred by optimal learning in humans might actually be outweighed by these computational costs. Related, we do not currently have a precise framework with which to precisely describe how ideal Bayesian learning, with its requirement for complex integrals, would be implemented neuronally, although there are ongoing efforts to establish underlying mechanisms, such as spiking neural networks (Deneve, 2007) and probabilistic population coding models (Sanger, 1996; Pouget et al., 2003; Ma et al., 2006; Yang and Shadlen, 2007; Beck et al., 2008; Ma and Jazayeri, 2014).

In addition, despite evidence that humans and other animals demonstrate considerable inter-individual variability in learning and action, even under carefully controlled experimental conditions (Gluck et al., 2002; Daunizeau et al., 2010b), many Bayesian models are agnostic to inter-individual variability. There have been some attempts to

construct Bayesian models capable of capturing inter-individual variability (Steyvers et al., 2009; Nassar et al., 2010). Nonetheless, the failure of traditional Bayesian learning theory to account for these individual differences remains a key problem for understanding (mal)adaptive learning and action in humans.

	Advantages	Disadvantages
Reinforcement learning models	• Simple update equations are computationally feasible • Intuitive framework for investigations of learning and adaptive behaviour • Non-normative nature allows for modelling of aberrant modes of learning	• Heuristic approach • Not grounded in probability theory • Often perform badly in real-world situations where environmental states and outcomes are not known in advance • Do not capture uncertainty
Bayesian learning models	• Capture the optimal method for integrating multiple sources of new information and prior knowledge • Grounded in probability theory • Capture uncertainty of beliefs	• Computationally demanding, making online learning difficult • Currently unclear how they would be implemented neuronally • Many models are agnostic to inter-individual variability

Table 3.1 The advantages and disadvantages of different learning models.

3.3 The Hierarchical Gaussian Filter

In 2011, Mathys et al. developed the Hierarchical Gaussian Filter (HGF) model as a generic hierarchical Bayesian framework for individual learning under the various forms of uncertainty inherent in dynamic environments (Mathys et al., 2011, 2014). The HGF takes inspiration from RL schemes and aims to overcome the limitations of Bayesian approaches, namely their computational complexity and failure to capture differences in learning across individuals. It uses variational Bayes under mean-field approximation to derive trial-wise update equations that are analytical and efficient, allowing for real-time learning. A novel approximation to the conditional probabilities over unknown quantities replaces the conventional Laplace approximation used in Bayesian schemes. The form of the update equations is similar to those used in RW learning, meaning the HGF provides a Bayesian analogue to, and has a natural interpretation in terms of, RL.

The HGF is an extension of a model proposed by Daunizeau et al., which quantifies the likelihood of an individual's observed behaviour based on Bayes-optimal inferences in probabilistic environments (Daunizeau et al., 2010b). It also draws inspiration from the

aforementioned work by Behrens et al. so as to capture alterations in learning under environmental volatility (Behrens et al., 2007). Briefly, the original instantiation of the HGF comprises a *perceptual model* that tracks an individual's learning of the underlying structure of the environment. The perceptual model has two components: a *generative model* and a *recognition model*. The generative model comprises a set of probabilistic assumptions that describe how sensory signals in the environment are generated. The recognition model captures the unobservable inference process made by an individual based on these sensory signals. It does this by performing (approximate) statistical inference on the observations of the actual sensory data and thus determining the probability distribution over variables in the generative model appropriate to those particular observations.

3.3.1 Perceptual model

The perceptual model makes it possible to quantify the inferences individuals make during an experimental learning task with known sensory signals, and to decompose the contributions of different forms of uncertainty to those inferences. In contradistinction to models that assume that individuals fashion the generative process to the task at hand (see the General Discussion for details), the HGF offers an inclusive scheme for explaining learning that generalises to a multitude of situations requiring inference about the state of the world.

3.3.1.1 Generative model

The HGF's generative model describes how the hidden environmental states of the world, x, generate sensory inputs, u, across three hierarchical levels (Figure 3.3). Hierarchical Bayesian models have proven powerful for explaining learning in volatile probabilistic environments (Behrens et al., 2007; den Ouden et al., 2010; Nassar et al., 2010; Wilson et al., 2013). In the case of the HGF, level 1 concerns trial-wise sensory outcomes, level 2 the probabilistic relationship between sensory outcomes and their predictive cues, and level 3 the volatility of this probabilistic relationship over time.

The original instantiation of the HGF models an environment in which trial-wise sensory outcomes are of a binary form. Therefore, at level 1, the environmental state x_1 at time t, denoted by $x^{(t)} \in \{0,1\}$, causes sensory input $u^{(t)}$. This might capture whether a visual stimulus is black or white, an olfactory stimulus is present or not present, or, in the case of the probabilistic learning task used in Chapter 6, whether an auditory stimulus is a vocalisation of the word "cow" or the word "pig". Accordingly, in what follows, the

likelihood model is assumed to take the following form (note that the time index *t* has been omitted here for simplicity):

$$p(u|x_1) = (u)^{x_1}(1-u)^{1-x_1}$$

Equation 3.5

Thus, $u = x_1$ for both $x_1 = 1$ and $x_1 = 0$ (where 1 = stimulus type A; 0 = stimulus type B), and vice versa. As such, x_1 captures the stimulus type. Knowing state x_1 allows for an accurate prediction of input u. x_1 is drawn from a Bernoulli distribution. The predicted probability of a particular sensory outcome is obtained by applying a sigmoid transformation to x_2:

$$p(x_1|x_2) = s(x_2)^{x_1}\big(1-s(x_2)\big)^{1-x_1} = \text{Bernoulli}\big(x_1; s(x_2)\big)$$

Equation 3.6

where s is the logistic sigmoid function:

$$s(x) = \frac{1}{1+\exp(-x)}$$

Equation 3.7

Thus, x_2 is mapped to the probability of x_1 such that $x_2 = 0$ means that $x_1 = 0$ and $x_1 = 1$ are equally probable.

At level 2, x_2 represents the probabilistic relationship between sensory cues and outcomes, in logit space. This could be, for instance, the conditional probability of an auditory outcome stimulus given an auditory cue (i.e., the cue:outcome contingency). It is an unbounded real parameter of the probability that $x_1 = 1$:

$$p\left(x_2^{(t)}\middle|x_2^{(t-1)}, x_3^{(t)}\right) = \mathsf{N}\left(x_2^{(t)};\, x_2^{(t-1)}, \exp\left(x_3^{(t)} + \omega\right)\right)$$

Equation 3.8

At level 3, x_3 represents the (log-)volatility of the environment:

$$p\left(x_3^{(t)}\middle|x_3^{(t-1)}, \vartheta\right) = \mathsf{N}\left(x_3^{(t)};\, x_3^{(t-1)}, \vartheta\right)$$

Equation 3.9

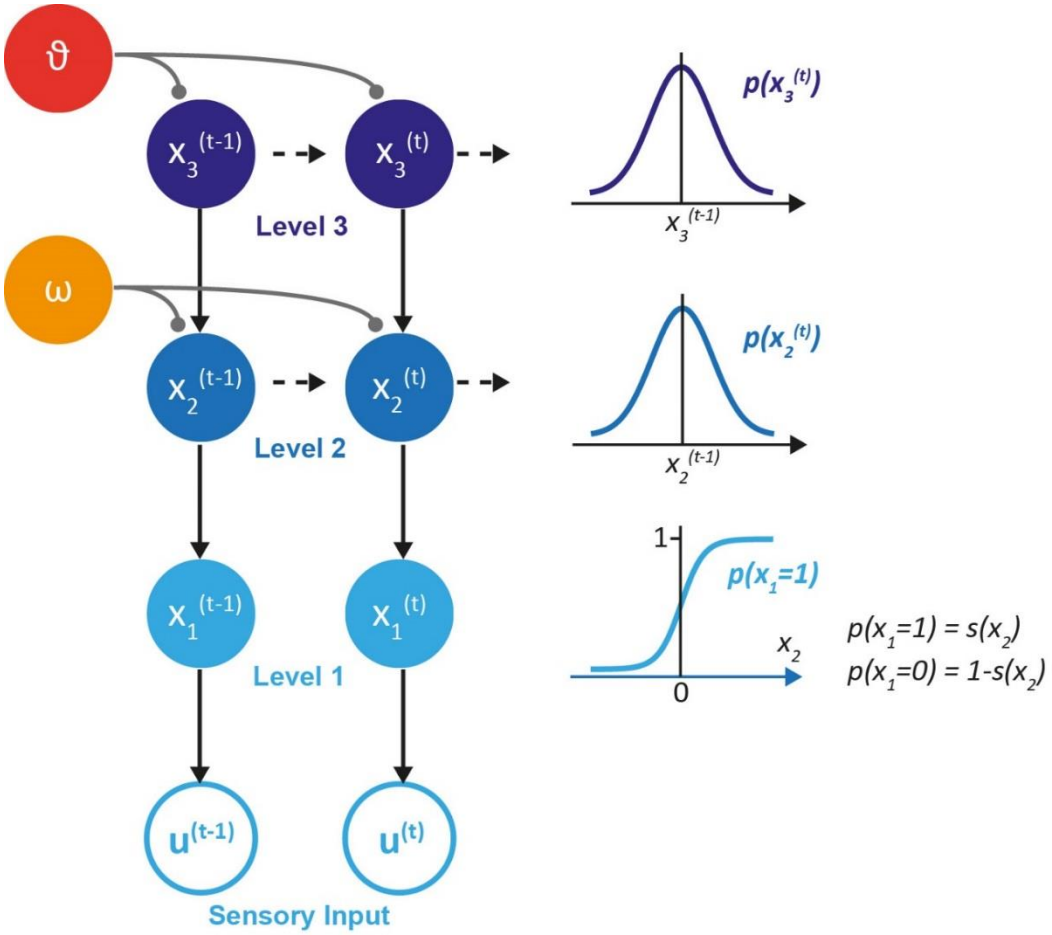

Figure 3.3 A schematic representation of the HGF's generative model. *Left:* $x_1^{(t)}$, $x_2^{(t)}$ and $x_3^{(t)}$ *are hidden states of the environment at time t.* $x_1^{(t)}$*represents the sensory outcome on the current trial,* $x_2^{(t)}$ *the probability of that outcome given a preceding sensory cue, and* $x_3^{(t)}$ *the volatility of the probabilistic relationship between cues and outcomes. The hidden states generate the sensory input at time t,* $u^{(t)}$*.The hidden states at levels 2 and 3 are dependent on their immediately preceding values,* $x_2^{(t-1)}$ *and* $x_3^{(t-1)}$*, and two participant-specific parameters, ϑ and ω, that link the hierarchical levels. Right: The probability of the hidden state at each level is determined by the variables and parameters at the next highest level. At level 1,* x_1 *determines the probability of the input* u*. At levels 2 and 3, the hidden states evolve as Gaussian random walks, with participant-specific step sizes. At level 2, the step size of the Gaussian random walk is captured by the variance exp(*x_3*+ω) of the conditional probability. At level 3, the step size is captured by ϑ. Figure adapted from Mathys et al. 2011.*

At levels 2 and 3, the states evolve in time as Gaussian random walks, with each walk's step size determined by the next highest level of the hierarchy. For the sake of generality,

the HGF makes no assumptions about the probabilities of $x_2^{(t)}$ or $x_3^{(t)}$, except that they may change over time as Gaussian random walks. This means that the values of $x_2^{(t)}$ and $x_3^{(t)}$ will be normally distributed around their values at the preceding time point, $x_2^{(t-1)}$ and $x_3^{(t-1)}$ respectively. Importantly, the HGF does not assume fixed "ideal" learning across individuals but rather contains participant-specific parameters that couple the hierarchical levels and allow for individual expression of (approximate) Bayes-optimal learning.

Specifically, the dispersion of the random walk at level 2 (i.e., the variance $\exp(x_3+\omega)$ of the conditional probability; Equation 3.8) is determined by both the participant-specific parameter ω and state x_3 (c.f. Behrens et al., 2007, 2008). ω is a constant component of volatility that captures how rapidly an individual generally updates their beliefs about probabilistic relationships within the environment. As such, a higher ω would lead an individual to update their beliefs about environmental contingencies more rapidly, resulting in a faster tonic learning rate. x_3 captures the environment's phasic volatility, which can vary over time. Introducing ω therefore allows for a participant-specific tonic component of volatility that scales independently of state x_3. The higher the volatility, the larger the step-size of the Gaussian walk at level 2. Note that the original instantiation of the HGF includes an additional participant-specific parameter, κ, which scales state x_3, and hence modulates coupling between levels 2 and 3. In all applications of the HGF in this thesis, κ was held constant at 1 (Vossel et al., 2014a, 2014b, 2015; de Berker et al., 2016).

At level 3, the step-size of the Gaussian walk is determined by the participant-specific parameter ϑ, which captures the volatility of state x_3 (Equation 3.9). ϑ therefore determines the speed of learning about volatility, i.e., the rate at which estimates of the environment's phasic volatility are updated. As such, ϑ encapsulates metavolatility, i.e., the rate at which volatility changes, with higher values leading to a Gaussian random walk with a larger step-size, and implying a belief in a more unstable world, in turn leading to a more variable learning rate at level 2. It is possible to add additional levels to the HGF's perceptual model, in which case the step-size of the Gaussian random walk would be determined by both ϑ and the state at level 4, x_4. Mirroring previous studies that have utilised the HGF to investigate hierarchical learning (Iglesias et al., 2013; Diaconescu et al., 2014, 2017; Hauser et al., 2014; Vossel et al., 2014a, 2014b, 2015; de Berker et al., 2016), a three-level perceptual model was applied in this thesis.

3.3.1.2 Recognition model

Based on their observations, x_1, individuals form and update beliefs about the true states represented at each level of the HGF. These beliefs are captured by a recognition model. Under this recognition model, individuals infer the approximate posterior distributions over the states at levels 2 and 3. More precisely, trial by trial, individuals update their beliefs about the true quantities at each level, which at levels 2 and 3 are modelled by Gaussian distributions with a mean (μ) and variance (σ), the latter reflecting the uncertainty of the estimate. The recognition model captures *irreducible uncertainty* arising from the inherent randomness of the probabilistic relationships between cues and outcomes at level 1, *estimation uncertainty* arising from an individual's incomplete knowledge of these probabilistic relationships at level 2, and *volatility uncertainty* arising from the instability of these relationships at level 3.

Sufficient statistics of the Gaussian approximations are computed at each time-point. The resulting update equations resemble RW learning (c.f. Equation 3.1) and take the form:

$$\text{prediction}^{(t)} = \text{prediction}^{(t-1)} + \text{learning rate x prediction error},$$

where *t* is the current time-point.

Given the aforementioned body of work suggesting that RL may be implemented neuronally, Mathys et al. postulate that this property means that the belief updates have an ecologically valid interpretation. Specifically, the HGF parameters that determine learning may relate to specific physiological processes, such as the neuromodulation of synaptic plasticity (Mathys et al., 2011). For instance, it has been hypothesised that dopamine, which regulates the plasticity of glutamatergic synapses (Gu, 2002), may encode the precision (i.e., the inverse variance, or inverse uncertainty) of PEs (Friston, 2009). As I will describe later, precision-weighting of PEs occurs within the HGF's computational framework. The HGF therefore offers a useful model-based approach to probing participant-specific computational and neurophysiological mechanisms of learning under uncertainty. Moreover, the dynamic learning rates that result from coupling the HGF's different levels allow for the adaptation necessary in volatile environments (Behrens et al., 2007; den Ouden et al., 2010).

Since the integrals arising in the recognition model are intractable, inference in the HGF is approximate. Variational Bayesian inversion determines the posterior distributions by maximising the log-model evidence, which corresponds to the negative surprise about the data, given a model. It is approximated by a lower bound, namely the negative free

energy (Beal, 2003; Friston and Stephan, 2007), building on the work of Friston et al. (Friston et al., 2006).

3.3.2 Response model

To link an individual's posterior beliefs, as provided by the recognition model, to his/her actions, a response model that provides a complete mechanistic mapping from experimental stimuli to observed behavioural responses is required (Daunizeau et al., 2010a, 2010b). This was a key focus of our novel instantiation of the HGF, which I describe in section 3.4.

3.3.3 The merits and shortcomings of the HGF

To summarise, the original instantiation HGF, with its constituent generative, recognition and response models, has the capacity to describe learning and action that is subjectively optimal in relation to an individual's prior beliefs and sensory input. It is possible for this learning to be objectively maladaptive. Importantly, this means that the HGF can capture variations in learning across healthy individuals, as well as aberrant belief updating in individuals with conditions such as schizophrenia (Adams et al., 2016). Thus, the HGF's approach to modelling learning and action may have potential clinical applications, including the development of diagnostic classifications of psychiatric spectrum disorders (Stephan et al., 2009a). The HGF has been successfully applied in several investigations of learning and action in volatile environments (Iglesias et al., 2013; Diaconescu et al., 2014, 2017; Hauser et al., 2014; Vossel et al., 2014a, 2014b, 2015; de Berker et al., 2016). In Chapter 6, I apply the original instantiation of the HGF to behavioural data recorded in individuals undertaking a probabilistic learning task with binary trial-wise outcomes.

As mentioned above, the HGF was designed as an inclusive scheme for explaining learning that generalises to different situations requiring inference about the state of the world. However, modelling an environment in which all trial-wise sensory outcomes are of a binary form is clearly not representative of real-world scenarios. Further, the HGF's perceptual model is not sufficient to elucidate the influence an individual's perceptual beliefs has on action execution.

As part of this thesis, I have worked on the development of a novel instantiation of the HGF. It features a perceptual model with the capacity to track an individual's learning about multiple outcome types and a response model that quantifies the influence of that individual's perceptual beliefs on their executed actions, specifically their RTs. Next, I describe how this novel instantiation of the HGF was applied to model individual learning

and action during the PSRTT utilised in Chapters 4 and 5. I also provide additional detail as to how individual learning and action is captured by the model.

3.4 Developing a novel instantiation of the HGF

To recap Chapter 2, the PSRTT required individuals to respond to the trial-wise presentation of one of four visual stimuli by pressing an appropriate button as quickly as possible. At any given time, the trial sequence was generated by one of eight transition matrices (TMs), which changed every 50 trials. In each case, there were 16 combinations that determined the probabilistic relationship between the stimuli presented on the current trial *t*, and the previous trial, *t-1*. The task design created contexts that participants could infer from their stimulus observations, allowing them to reduce their uncertainty about sensory events before they occurred (Harrison et al., 2006). Nonetheless, the probabilistic nature of these contexts also produced unexpected stimulus outcomes, i.e., sensory PEs. For fast and accurate responses, participants had to track *irreducible uncertainty* arising from the inherent randomness of the probabilistic transitions between consecutive stimuli; *estimation uncertainty* arising from their imperfect knowledge of the probabilistic relationships governing stimulus transition contingencies within contexts; and *volatility uncertainty* maintained by the unsignalled contextual instability.

The novel instantiation of the HGF has a focus on transition matrices and includes two components: a three-level perceptual model and a response model (Figure 3.4). The relevant Matlab code has been incorporated into the HGF Toolbox, which is available for download from http://www.translationalneuromodeling.org/tapas/. The perceptual model encompasses a generative model that describes how stimulus transitions were generated and a recognition model that captures an individual's unobservable beliefs about these transitions. Since beliefs modulate behaviour, it is possible to reverse engineer an individual's observed actions (here trial-wise log(RTs)) to infer their beliefs. A response model was developed to predict normally-distributed log(RTs) from parameters in the recognition model (Daunizeau et al., 2010a).

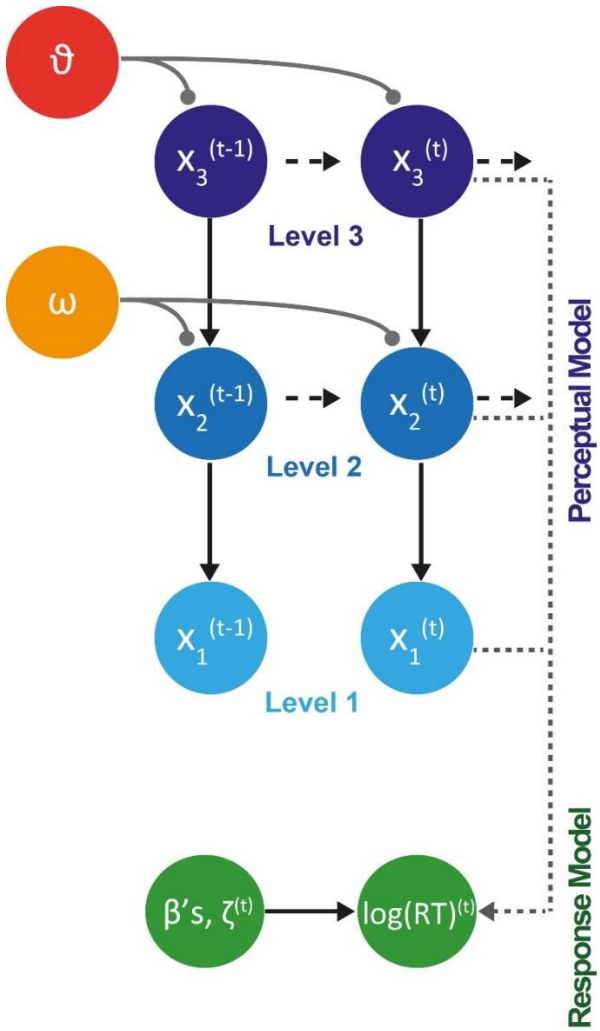

Figure 3.4 A novel instantiation of the HGF. *The perceptual model tracks an individual's learning of the PSRTT's structure across three levels. State* $\boldsymbol{x}_1$ *represents trial-wise stimulus transitions from one stimulus to the next,* $\boldsymbol{x}_2$ *the transition contingencies, and* x_3 *the phasic volatility of these contingencies, where t is the current trial number. Participants hold and update beliefs about the true quantities at each level, with a mean μ and a variance σ. ϑ and ω are participant-specific parameters that couple the levels and determine the respective speed of belief updating about phasic volatility and transition contingencies. The response model describes the mapping from a participant's trial-wise beliefs onto their observed log(RT) responses.*

3.4.1 Perceptual model

The perceptual model is a variant of the HGF as introduced by Mathys et al. (Mathys et al., 2011, 2014). It comprises a generative model and a recognition model, and thus tracks a participant's learning of the task's structure in a similar way to previous studies using the HGF (Iglesias et al., 2013; Diaconescu et al., 2014, 2017; Hauser et al., 2014; Vossel et al., 2014a, 2014b, 2015; de Berker et al., 2016).

3.4.1.1 Generative model

Unlike previous applications of the HGF, in the novel instantiation the sensory data are observed transitions between stimuli that arise from a sequence of environmental states ($\boldsymbol{x}_1$), where bold font is used to indicate a matrix. In the PSRTT (Figure 2.1), the jkth element of $\boldsymbol{x}_1$ is the transition from stimulus k to stimulus j, the probability of which participants must learn to perform the task well. There are 16 possible transitions induced by the trial-wise presentation of one of four visual stimuli, meaning that $\boldsymbol{x}_1$ is a four-by-four matrix. On each trial, an individual observes a sample from one column of the transition matrix. Therefore, the current transition in the corresponding column of $\boldsymbol{x}_1$ is 1, with all other elements in that column equal to zero.

The generative model has two further levels above $\boldsymbol{x}_1$. Level 2 is a four-by-four matrix $\boldsymbol{x}_2$ of real numbers governing the transition contingencies. These undergo random walks with increments that are independent of each other. At level 3, x_3 sets the variance of those random walks, and so the rate of change (or volatility) of the elements of $\boldsymbol{x}_2$. Since all elements are assumed to experience the same volatility (c.f. Mathys et al. 2011; 2014), x_3 is a scalar. Collectively, $\boldsymbol{x}_2$ and x_3 capture stimulus transitions and their changes over time (albeit represented heuristically as a continuous random walk in logit space with a bijective mapping to the probability of specific discrete changes). More specifically, a sample of $\boldsymbol{x}_1$ is generated by applying a logistic sigmoid transformation to the column of $\boldsymbol{x}_2$ associated with the stimulus that was previously shown to generate a probability distribution over the four possible next stimuli. A sample is then drawn from that distribution.

Thus, level 1 of the HGF represents a sequence of environmental states $\boldsymbol{x}_1$ (here the presentation of one of four stimuli). Level 2 represents the transition contingency $\boldsymbol{x}_2$ (i.e., the conditional probability, in logit space, of the stimulus on trial *t* given the stimulus presented on trial *t-1*). x_3 represents the phasic volatility. The hidden states at levels 2 and 3 are assumed to evolve as a Gaussian random walk, such that their variance depends on the state at the level above:

$$p\left(x_{1,jk}\middle|x_{2,jk}\right) = s\left(x_{2,jk}\right)^{x_{1,jk}}\left(1 - s\left(x_{2,jk}\right)\right)^{1-x_{1,jk}} = \text{Bernoulli}\left(x_{1,jk}; s\left(x_{2,jk}\right)\right)$$

Equation 3.10

$$p\left(x_{2,jk}^{(t)}\middle|x_{2,jk}^{(t-1)}, x_3^{(t)}\right) = \mathsf{N}\left(x_{2,jk}^{(t)};\, x_{2,jk}^{(t-1)}, \exp\left(x_3^{(t)} + \omega\right)\right)$$

Equation 3.11

$$p\left(x_3^{(t)}\middle|x_3^{(t-1)}, \vartheta\right) = \mathsf{N}\left(x_3^{(t)};\, x_3^{(t-1)}, \vartheta\right)$$

Equation 3.12

where $x_{1,jk}$ and $x_{2,jk}$ (with j,k=1,…,4) are the elements of the level 1 transition matrix $\boldsymbol{x}_1$ and of the level 2 matrix $\boldsymbol{x}_2$ respectively, and s is the logistic sigmoid function previously defined in Equation 3.7.

3.4.1.2 Recognition model

The recognition model takes observations of $\boldsymbol{x}_1$ and infers approximate posterior distributions over the values of $\boldsymbol{x}_2$ and x_3. This amounts to a variant of predictive coding in which beliefs are dynamically updated across the levels via PEs that are weighted by their salience, or expected precision (equivalent to inverse variance, or uncertainty). Estimates of stimulus transition contingencies correspond to the posterior distribution over $\boldsymbol{x}_2$ and are updated by PEs about stimulus occurrences. Estimates of environmental volatility, i.e., the posterior distribution over x_3, are updated in proportion to PEs about the transition contingencies. Thus, the effective learning rate is influenced by uncertainty about current beliefs and environmental instability.

As in the original instantiation of the HGF, trial by trial, participants update their beliefs about the true quantities at each level, which at levels 2 and 3 are modelled by Gaussian distributions with a mean (μ) and variance (σ), the latter reflecting the uncertainty of the estimate. Precision ($\hat{\pi}$) of the prediction is equal to the inverse variance ($1/\hat{\sigma}$), where the hat denotes the participant's predicted estimate before seeing the stimulus outcome on each trial. At level 1, when the elements of $\boldsymbol{x}_2$ are each transformed by the logistic sigmoid to produce probabilities $\hat{\boldsymbol{x}}_1$, there is *irreducible uncertainty* (the participant's estimate of which is captured by $\hat{\boldsymbol{\mu}}_1$). Irreducible uncertainty, which gets its name since it is undiminished by learning (Payzan-LeNestour and Bossaerts, 2011), arises from any probabilistic relationship, and is closely related to entropy, with an inverted-U relationship to probability that peaks at p=0.5. The quantity gives rise to sensory PE ($\boldsymbol{\delta}_1$) following the presentation of an unexpected, or *surprising*, stimulus that would require a participant to respond against their expectation:

$$\delta_{1,jk}^{(t)} = \mu_{1,jk}^{(t)} - \hat{\mu}_{1,jk}^{(t)}$$

Equation 3.13

where the prediction $\hat{\boldsymbol{\mu}}_1^{(t)}$ about stimulus outcome results from a sigmoidal transformation of the previous belief about the stimulus transition contingency $\boldsymbol{\mu}_2^{(t-1)}$:

$$\hat{\mu}_{1,jk}^{(t)} = s(\mu_{2,jk}^{(t-1)})$$

Equation 3.14

Note that column-wise normalisation of $\hat{\boldsymbol{x}}_1$ is not enforced (i.e., the columns do not necessarily add up to one, as they would have to in order to represent a probability distribution over mutually exclusive events). Ensuring that the probabilities sum to one would arguably require a sort of certainty about the stimuli that participants do not necessarily have when performing the behavioural task; for instance, it would require precise *a priori* knowledge that each and every trial will present exactly one of four stimuli and that there is no possibility of novel stimuli occurring during the experiment. In practice, the statistics governing the sensory events that occur in the PSRTT ensure that column sums of participants' $\hat{\boldsymbol{\mu}}_1$ estimates never stray far from unity.

At level 2, $\boldsymbol{\sigma}_2$ which is informational in origin, represents *estimation uncertainty* about the true probabilistic relationships governing stimulus transitions, giving rise to a more abstract contingency PE ($\boldsymbol{\delta}_2$). At level 3, *volatility uncertainty* arises from the environment's volatility, i.e. how quickly the transition contingencies are changing. This is in contrast to σ_3, which represents the uncertainty about the volatility.

Generally, at any level i of the hierarchy, the update of the belief on trial *t* (i.e., posterior mean $\mu_i^{(t)}$ of the state x_i) is proportional to the precision-weighted PE, $\varepsilon_i^{(t)}$. This weighted PE is the product of the upward-propagating PE, $\delta_{i-1}^{(t)}$, and a precision ratio, $\psi_i^{(t)}$, capturing the uncertainty about input from the level below relative to the uncertainty about the state of the level being updated (Iglesias et al., 2013). A general and didactically useful form of this precision-weighted PE (with subtle differences below level 3; see Mathys et al. 2014) is:

$$\Delta\mu_i^{(t)} \propto \varepsilon_i^{(t)} = \psi_i^{(t)}\,\delta_{i-1}^{(t)}$$

Equation 3.15

where

$$\psi_i^{(t)} = \frac{\hat{\pi}_{i-1}^{(t)}}{\pi_i^{(t)}}$$

Equation 3.16

Thus, precision-weighted sensory PE ($\boldsymbol{\varepsilon}_2$) about stimulus outcome is weighted by uncertainty at levels 1 and 2 and serves to update the belief about $\boldsymbol{x}_2$ (the stimulus transition contingency in logit space):

$$\mu_{2,jk}^{(t)} - \mu_{2,jk}^{(t-1)} = \psi_{2,jk}^{(t)} \delta_{1,jk}^{(t)}$$

$$= \varepsilon_{2,jk}^{(t)}$$

Equation 3.17

At level 3, the update of the belief about x_3 (phasic environmental (log-)volatility) is proportional to the precision-weighted contingency PE $\boldsymbol{\varepsilon}_3$, which captures uncertainty at levels 2 and 3:

$$\mu_3^{(t)} - \mu_3^{(t-1)} \propto \psi_{3,jk}^{(t)} \delta_{2,jk}^{(t)}$$

$$= \frac{\hat{\pi}_{2,jk}^{(t)}}{\pi_3^{(t)}} \delta_{2,jk}^{(t)}$$

$$= \varepsilon_{3,jk}^{(t)}$$

Equation 3.18

Here, the PE concerns the volatility of the stimulus transition contingency, or more precisely, the variance ratio of its estimates (in logit space) after and before observing the sensory input, respectively:

$$\delta_{2,jk}^{(t)} = \frac{\sigma_{2,jk}^{(t)} + (\mu_{2,jk}^{(t)} - \mu_{2,jk}^{(t-1)})^2}{\sigma_{2,jk}^{(t-1)} + e^{\mu_3{}^{(t-1)+\omega}}} - 1$$

Equation 3.19

Importantly, like the original instantiation, the novel perceptual model includes two participant-specific parameters that couple the hierarchical levels and allow for individual expression of approximate Bayes-optimal learning (Figure 3.4). The first of these parameters is ϑ, which determines the speed of learning about the volatility of the environment, i.e. the rate at which estimates of trial-wise phasic volatility (μ_3) are updated

(Equation 3.12). The second, ω, is a constant component of the learning rate at level 2 that captures a tonic learning rate about the stimulus transition contingencies (Equation 3.11).

The punctate change-points contained in the true generative process are detected implicitly by the HGF via spikes in the precision weights. At levels 2 and 3, $\alpha_i^{(t)}$ is proportional to the precision ratio, $\psi_i^{(t)}$, defined in Equation 3.16. At level 1, the learning rate $\boldsymbol{\alpha_1}$ is simply defined as the update divided by the prediction error:

$$\boldsymbol{\alpha}_1^{(t)} \propto \frac{\boldsymbol{\mu}_2^{(t)} - \hat{\boldsymbol{\mu}}_1^{(t)}}{\boldsymbol{\delta}_1^{(t)}}$$

Equation 3.20

As I will demonstrate in Chapter 4, the HGF implicitly captures punctate change-points in the PSRTT's generative process as an increase in learning rate, α_1, following a true change in context (Figure 3.5).

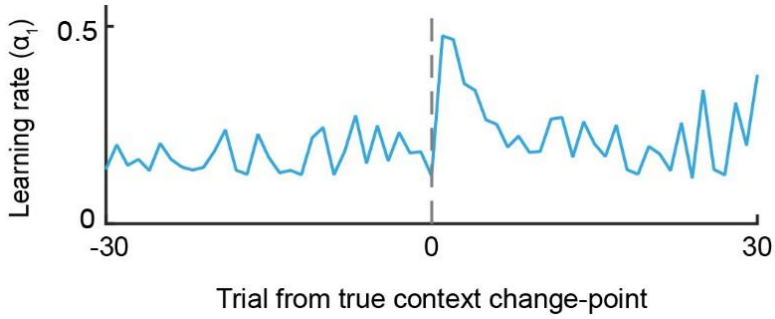

Figure 3.5 Example learning rate (α_1) trajectory. *Increases in α_1 are observed following a true change in context. As such, learning rate at level 1 of the HGF implicitly captures punctate change-points contained in the PSRTT's generative process. Data are mean α_1 for trials undertaken by Placebo participants in Chapter 4. For further details, see Figure 4.5.*

3.4.2 Response model

As described above, in addition to the perceptual (generative + recognition) model, the novel instantiation of the HGF features a response model. Its purpose is to link estimates from the recognition component of the perceptual model to an individual's actions during the PSRTT. A response model offers an important extension to the perceptual model by linking modulations of action execution to perceptual beliefs. Alternative response models have been added to the HGF previously (Vossel et al., 2014a, 2014b, 2015). The response model developed and applied in the present work describes the mapping from

a participant's trial-wise beliefs, as provided by the perceptual model, onto his/her observed responses, log(RTs).

I reasoned that there are several variables that could influence this mapping, and thus trial-wise log(RT). Therefore, I constructed and formally compared a range of response models using random effects Bayesian model selection (Stephan et al., 2009; Rigoux et al., 2014) and associated techniques for assessing differences in model frequencies across groups, as implemented in the VBA toolbox (Daunizeau et al., 2014). Further details of this model comparison are provided in Chapter 4.

As we will see in Chapter 4, the winning response model (Equation 3.21) was a linear function prescribing that trial-wise log(RT) is determined by a constant component of RT (β_0) and estimates arising from each level of the perceptual model: sensory PE (δ_1) arising at level 1, precision-weighted contingency PE (ε_3) arising at level 2, and trial-wise phasic volatility (μ_3) arising at level 3. Additionally, as observed in previous work using similar RT tasks (Rabbitt, 1966; Botvinick et al., 2001; Gehring and Fencsik, 2001; Cavanagh et al., 2014), there was evidence of post-error slowing in the PSRTT, i.e., participants were slower to respond on a trial that followed an incorrect response. ζ is a noise term.

$$\log(\mathrm{RT})^{(t)} = \beta_0 + \beta_1(\delta_1^{(t)}) + \beta_2(\varepsilon_3^{(t)}) + \beta_3(\mu_3^{(t)}) + \beta_4(\mathrm{PostError}^{(t)}) + \zeta^{(t)}$$

Equation 3.21

While the perceptual model assumes that participants update their beliefs according to the stimulus presented on each trial, the response model incorporates correct trials only.

3.4.3 Model fitting

For each participant, individual maximum *a posteriori* estimates for perceptual and response model parameters were jointly obtained using the Broyden-Fletcher-Goldfarb-Shanno algorithm as implemented in the HGF Toolbox. Where priors were required, they were defined by inverting the perceptual model in isolation, given the known stimulus sequence (using the function 'tapas_bayes_optimal_whatworld_config' contained in the TAPAS Toolbox), under suitably uninformative priors. The resulting posterior estimates were then used to define the priors for the subsequent inversion of the full model given the behavioural data. In other words, the prior means in the empirical data analysis corresponded to those parameter values for which the stimulus sequence would generate minimal surprise (in an observer with the aforementioned uninformative priors).

4 Pharmacological fingerprints of uncertainty

*This chapter is based on work presented in **Marshall L**, Mathys C, Ruge D, de Berker AO, Dayan P, Stephan KE & Bestmann S. (2016) Pharmacological fingerprints of contextual uncertainty. PLOS Biology. 14(11): e1002575.*

4.1 Abstract

Successful interaction with the environment requires flexible updating of our beliefs about the world. By estimating the likelihood of future events, it is possible to prepare appropriate actions in advance and execute fast, accurate motor responses. According to theoretical proposals, individuals track the variability arising from changing environments by computing various forms of uncertainty. Several neuromodulators have been linked to uncertainty signalling, but comprehensive empirical characterisation of their relative contributions to perceptual belief updating, and to the selection of motor responses, is lacking. In this chapter, I assess the roles of noradrenaline (NA), acetylcholine (ACh) and dopamine (DA) within a single, unified computational framework of uncertainty. Using pharmacological interventions in a sample of 128 healthy human participants and a hierarchical Bayesian learning model, I characterise the influences of noradrenergic, cholinergic and dopaminergic receptor antagonism on individual computations of uncertainty during a probabilistic serial reaction time task. I propose that NA influences learning of uncertain events arising from unexpected changes in the environment, while ACh balances attribution of uncertainty to chance fluctuations within environmental contexts or to gross environmental violations following a contextual switch. In contrast, DA supports the use of uncertainty representations to engender fast, adaptive motor responses.

4.2 Introduction

Adaptive performance in dynamic environments depends on our ability to represent and manipulate internal estimates of the world's statistical structure (Conant and Ashby, 1970; Körding and Wolpert, 2004; Yu and Dayan, 2005; Behrens et al., 2007). By tracking the environment's underlying regularities, an individual can learn the causes of their sensory input and thus the likelihood that a particular event will occur. In turn, this permits anticipatory action preparation and the rapid execution of responses (Bestmann et al., 2008).

However, the environment's richly complicated sources of noise and latent structure present us with various forms of uncertainty. In Chapter 1, I introduced three distinct forms. First, *irreducible uncertainty* captures the randomness inherent in any complex environment and is undiminished by learning. Second, *estimation uncertainty* arises from an individual's incomplete knowledge of the probabilistic relationships *within* the current environmental context. Third, *volatility uncertainty* arises from our beliefs about the stability of the environment, and thus how quickly probabilistic relationships are changing *between* contexts. Optimal learning, prediction and anticipatory action preparation require that these sources of uncertainty are taken into account (Ma and Jazayeri, 2014; Meyniel et al., 2015; Pouget et al., 2016).

4.2.1 The brain computes different forms of uncertainty

To recap Chapter 1, multiple lines of theoretical, behavioural and neurobiological evidence suggest that the brain indeed computes estimates of uncertainty relating to the environment's sensory events, contextual associations and their changes over time (Averbeck et al., 2006; Ma et al., 2006; Behrens et al., 2007; den Ouden et al., 2010; Fiser et al., 2010; Mathys et al., 2011, 2014; Payzan-LeNestour and Bossaerts, 2011; Bach and Dolan, 2012; Bland and Schaefer, 2012; Friston et al., 2012; Iglesias et al., 2013; Payzan-LeNestour et al., 2013; Vossel et al., 2014a, 2014b; de Berker et al., 2016; Diaconescu et al., 2017). It has been proposed that, with their broad distribution and extensive connectivity, the brain's neuromodulatory networks are well-placed to facilitate the widespread changes in neuronal gain required to modulate the relative impact of top-down prior expectations and bottom-up sensory evidence in light of uncertainty (Berridge and Waterhouse, 2003; Warren et al., 2016). In accordance with this notion, ACh and NA are known to enhance bottom-up, feedforward thalamocortical transmission of sensory information relative to top-down, intracortical and feedback processing (Hasselmo et al., 1996; Gil et al., 1997; Kimura et al., 1999; Kobayashi et al., 2000; Yu

and Dayan, 2002, 2005; Hasselmo and McGaughy, 2004; Sarter et al., 2005; Dayan and Yu, 2006a; Deco and Thiele, 2011; Moran et al., 2013), in turn promoting learning about the current environmental context (Yu and Dayan, 2003).

4.2.1.1 A proposed role for acetylcholine under estimation uncertainty

The application of two types of behavioural paradigm has offered more detailed insight into the relative roles played by NA and ACh in learning under uncertainty. Accordingly, ACh is thought to support learning *within* stable environmental contexts defined by particular rules. Here uncertainty arises from ignorance about, and the unreliability of, probabilistic relationships within the environment that predict upcoming sensory events. The learning of these relationships is modulated by pharmacological (Witte et al., 1997; Phillips et al., 2000a), surgical (Voytko et al., 1994; Chiba et al., 1999), and neurodegenerative (Parasuraman et al., 1992) manipulations of ACh. Moreover, activity in the human cholinergic basal forebrain has been shown to reflect an individual's estimation uncertainty about contextual probabilistic relationships (Iglesias et al., 2013; Diaconescu et al., 2017), while pharmacological cholinergic stimulation under the drug galantamine increases the rate at which humans learn probabilistic relationships under estimation uncertainty (Vossel et al., 2014a). Together, these findings support the notion that ACh enhances learning accorded to stimuli with uncertain predictive consequences (Bucci et al., 1998) by boosting the contribution of bottom-up sensory processing relative to top-down prior expectations (Yu and Dayan, 2005).

4.2.1.2 A proposed role for noradrenaline under environmental volatility

While NA plays no consistent role in probabilistic learning within contexts (Clark et al., 1989; Witte and Marrocco, 1997), it is thought to offer an interrupt signal when volatility uncertainty arises *between* contexts (Clark et al., 1989; Arnsten and Contant, 1992; Smith et al., 1992; Coull et al., 1995; Witte and Marrocco, 1997; Bouret and Sara, 2005; Dayan and Yu, 2006b). Learning to make accurate predictions from the strongly unexpected observations that follow a contextual switch necessitates heightened sensory vigilance and a disregard for outdated top-down expectations. NA, with its broad neural network capable of triggering multiple, simultaneous changes across the brain (Bouret and Sara, 2004), is well-placed to rapidly coordinate this process. Indeed, neurons in the locus coeruleus (LC), the primary source of cortical NA, show strong responses to unexpected environmental changes (Sara and Segal, 1991; Aston-Jones et al., 1997). Pharmacologically upregulating NA accelerates the detection of unexpected switches in the predictive properties of sensory stimuli (Devauges and Sara, 1990), while noradrenergic deafferentation of rat medial frontal cortex impairs

behavioural adaptation to contextual switches (McGaughy et al., 2008). Moreover, BOLD activity in the human LC has been shown to dynamically track volatility uncertainty (Payzan-LeNestour et al., 2013). Furthermore, pupil dilation – which is influenced by (nor)adrenergic afferents (Joshi et al., 2016) – correlates with unexpected changes in probabilistic context (Preuschoff et al., 2011; Nassar et al., 2012).

4.2.2 Motor responses are sensitive to uncertainty

Thus, uncertainty representations existing both within and between environmental contexts are crucial for optimal predictions about the probability of future events. While good predictions facilitate anticipatory preparation of appropriate motor responses (Bestmann et al., 2008), they are not sufficient for adaptive performance in dynamic environments. An additional mechanism is required to modify action selection based on one's own beliefs about the latent changes in the environment and/or the occurrence of unexpected events. Indeed, when an unexpected event occurs, humans are capable of engaging resources to inhibit a prepared response and replace it with an alternative (Hikosaka and Isoda, 2010; Isoda and Hikosaka, 2011), albeit at the expense of a prolonged reaction time (RT) (Galea et al., 2012; Bestmann et al., 2014).

4.2.2.1 A proposed role for dopamine in response modulation

There is considerable evidence linking DA to flexible behaviour (Cools et al., 2001a, 2009; Stelzel et al., 2010, 2013; van Holstein et al., 2011). Dopaminergic deficits due to Parkinson's disease are associated with specific flexibility impairments in both motor (Cools et al., 1984; Galea et al., 2012) and cognitive domains (Beatty and Monson, 1990; Cools et al., 2001a), with performance restored by dopaminergic medication (Cools et al., 2001b; Galea et al., 2012). In healthy individuals, pharmacological DA depletions impair adaptive reactions to unexpected events occurring within a broadly predictable context (Bestmann et al., 2014). However, it remains unclear whether DA supports accurate response selection by facilitating perceptual belief updating (Iglesias et al., 2013), or by modulating the sensitivity of response selection to perceptual beliefs.

4.2.3 A unified framework of uncertainty

In sum, while physiological, pharmacological, behavioural and theoretical work has suggested separable neuromodulatory involvement in different uncertainty computations, attempts to characterise the relative roles of NA, ACh and DA within a single computational scheme are lacking. Of note, since the conception of the work reported in this thesis, Varazzani et al. have contrasted the roles of NA and DA in motivation (Varazzani et al., 2015) and Brown et al. have assessed the impact of

pharmacological NA and ACh manipulations on orienting responses to novel stimuli (Brown et al., 2015). Nonetheless, there has been no direct investigation of the relative contributions of NA, ACh and DA to human learning and response modulation within a unified computational framework of uncertainty.

In this chapter, I employ the probabilistic serial RT task (PSRTT) introduced in Chapter 2 and the novel instantiation of the Hierarchical Gaussian Filter (HGF) model (Mathys et al., 2011, 2014) developed in Chapter 3 to characterise human learning and response modulation in dynamic, probabilistic environments and under pharmacological NA, ACh and DA interventions. To recap, the HGF's three-level *perceptual model* captures an individual's mapping from environmental causes to sensory inputs, while the *response model* maps those inferred environmental causes to observed RT responses (Daunizeau et al., 2010a). Thus, I sought to disentangle the effects of the three pharmacological manipulations on participant-specific perceptual belief updating under irreducible, estimation and volatility uncertainty from those effects on the sensitivity of motor responses to perceptual estimates.

4.3 Methods

4.3.1 Participants

128 healthy participants (56 male, aged 18-38 years, 119 right-handed) with normal or corrected-to-normal vision took part in this study after giving written informed consent. The experimental protocol was approved by the UCL Research Ethics Committee. The following exclusion criteria applied: history of neurological or psychiatric disease, intake of medication (other than contraceptives), self-reported smoking, self-reported recreational drug use, and current participation in other pharmacological studies. Following a screening interview to rule out intolerances or contraindications, the study clinician assigned participants pseudorandomly (i.e., ensuring a balanced distribution of gender, age and body weight) to receive a NA, ACh or DA antagonist, or a placebo. The experimenter (L.M.) was blind to the drug conditions.

4.3.2 General procedure

A double-blind, between-subjects design was employed. Each participant attended one experimental session during which they received a single, oral dose of one of the following: 1mg prazosin (α1-arenoceptor antagonist; NA- group), 6mg biperiden (M1-receptor antagonist; ACh- group), 2.5mg haloperidol (D1/D2-receptor antagonist; DA-group), or a placebo. Doses were selected in line with previous studies showing clear

behavioural and neurophysiological effects (Ziemann et al., 1997; Meintzschel and Ziemann, 2006; Korchounov and Ziemann, 2011; Bestmann et al., 2014). On arrival, participants completed computerised versions of the Digit Span test, Barratt Impulsiveness Scale (BIS-11) (Patton et al., 1995), Doman-Specific Risk-Taking (DOSPERT) Scale (Blais and Weber, 2006) and Cognitive Failures Questionnaire (CFQ) (Broadbent et al., 1982). Participants also self-reported their baseline mood (alertness, calmness and contentedness) with visual analogue scales (VAS) (Bond and Lader, 1974), and had their baseline heart rate (HR) and blood pressure (BP) measured. To assess any subjective and/or physiological drug effects, the VAS, HR and BP measurements were repeated before participants started the PSRTT and again once they completed it. Please refer to Chapter 2 (section 2.3.3) for details about the psychometric and subjective measures.

Two different drug administration times were used to match peak plasma concentration across drugs, based on previous pharmacokinetic data. To ensure that participants undertook the RT task when the drug was at its most active, haloperidol was administered two hours in advance (Time A; Figure 4.1A), while prazosin and biperiden were administered 1.5 hours before the main experimental session (Ziemann et al., 1997; Meintzschel and Ziemann, 2006; Korchounov and Ziemann, 2011). A random 50% of participants from the Placebo group were administered a placebo tablet at the first timepoint, and the other 50% at the second timepoint. The study clinician administered the drug or placebo while the experimenter was away from the testing room. Participants were asked not to eat for at least one hour before the first drug administration time.

4.3.3 Probabilistic serial reaction time task

Participants sat facing a computer screen positioned approximately 60cm away. They were instructed to rest their left and right index and middle fingers on the four buttons of a custom-made button box placed in front of them, and to maintain this position throughout the task. On each trial, participants were required to respond to the presentation of one of four visual stimuli by making a speeded button-press before the end of a 1200ms intertrial interval (ITI) (Figure 4.1B). Each stimulus was associated with one particular button. The stimulus-response mappings remained consistent within an experimental session but were counterbalanced across participants.

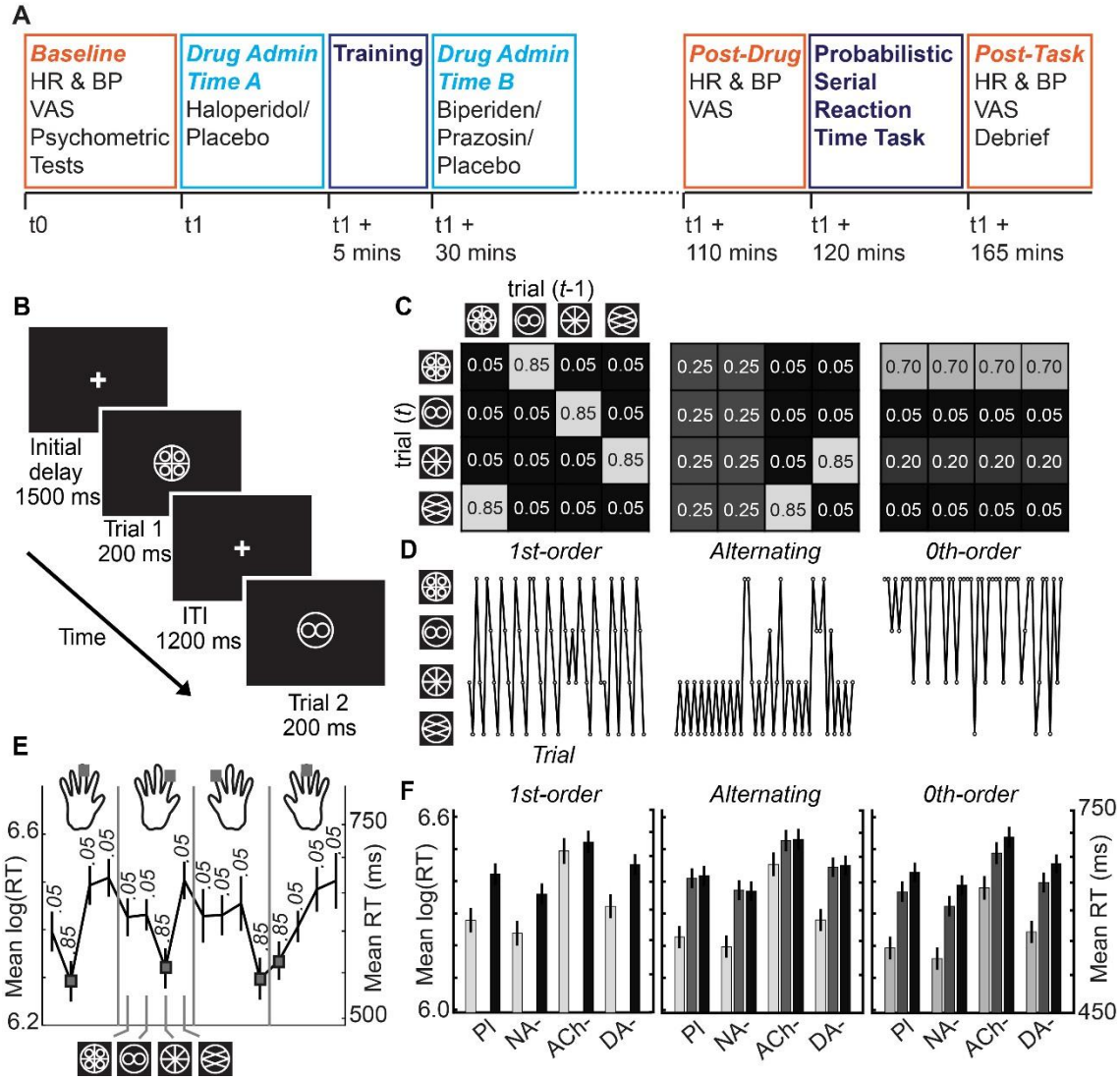

Figure 4.1 Task design. *(A) Timeline for each experimental session. At baseline, participants had their heart rate (HR) and blood pressure (BP) measured, self-reported their alertness, contentedness and calmness via visual analogue scales (VAS) (Bond and Lader, 1974), and undertook a battery of psychometric tests to assess working memory, impulsivity, risk-taking and distractibility. HR, BP and VAS measures were repeated before and after completing the behavioural task. Due to different times-to-peak plasma concentration across drugs, two different drug administration times (Time A and Time B) were used so that participants undertook the behavioural task when the drugs were at their most active. 50% of participants in the Placebo group received a placebo tablet at Time A and the other 50% at Time B. (B) Trial sequence. A trial began with the presentation of a central white fixation cross against a black background. After an initial delay of 1500ms at the start of each block, one of four visual stimuli was presented for 200ms. Participants were required to make a speeded button-press response before the end of a 1200ms intertrial interval (ITI). (C) Stimulus transitions were generated by one*

of eight different transition matrices (TMs), which changed every 50 trials without explicit indication to the participant. These TMs comprised two different 1st-order sequences, two alternating sequences, and four 0th-order sequences, each of which occurred three times in a pseudorandom order across 1200 trials. The overall probability of each stimulus was equal across the 1200 trials. For full details, see Figure 2.1. (D) Example trial sequences generated from the three example TMs in B. (E) By tracking the transition probabilities, subjects could learn to predict high probability events and prepare to make the correct button-press accordingly. Faster responses were observed for predictable stimuli compared to unexpected stimuli. Here Placebo group log(RTs) (mean ± SEM) are depicted for each of the 16 possible combinations between consecutive stimuli for the 1st-order sequence shown in C. Grey boxes indicate stimulus combinations with a high transition probability. (F) Indeed, across all types of TM, responses were faster for stimuli with higher transition probabilities (mean ± SEM).

4.3.3.1 Training

Each participant acquired the stimulus-response mappings for their session during a training block in which they received visual error feedback after each trial. The training session comprised at least 100 trials and did not finish until the participant had reached a minimum performance criterion of 85% accuracy on the last 20 trials. Participants were then given 40 practice trials, in which the stimuli were presented in a random order and without error feedback, to familiarise them with the timings of the main experiment. An additional refresher block, consisting of at least 26 trials with error feedback, was completed immediately before the main experiment. Again, participants had to achieve 85% accuracy in the last 20 trials to proceed. On average, participants reached this criterion in 28.1 ± 1.1 trials, indicating adequate learning and retention of the mappings. There was no difference in the number of refresher trials required between groups ($F_{3,120}$=1.17, p=0.324).

4.3.3.2 Task design

Each participant performed 1200 trials of the probabilistic RT task. Figure 4.1B shows an example trial sequence. Anticipatory responses (<80ms RT) were recorded as incorrect. At any given time, the trial sequence was generated by one of eight transition matrices (TMs), which changed every 50 trials without explicit indication to the participant. In each TM, there were 16 combinations that determined the probabilistic relationship between the stimulus presented on the current trial, *t*, and the stimulus presented on the previous trial, *t-1*. Three types of TM were utilised: two 1st-order sequences, two Alternating sequences, and four 0th-order sequences (see Figure 4.1C

and Figure 2.1 for further details). Trials were drawn from each TM three times. The order of TMs was pseudorandom, with no consecutive repeats. Importantly, the overall probability of each stimulus was equal across the 1200 trials.

The different TMs created contexts that the participants could infer from stimulus observations, allowing them to reduce their uncertainty about events before they occurred (Harrison et al., 2006). Nonetheless, the probabilistic nature of these contexts also produced unexpected stimulus outcomes, i.e. a sensory prediction error (PE). For fast and accurate responses, participants had to track three forms of uncertainty: *irreducible uncertainty* arising from the inherent randomness of the probabilistic transitions between consecutive stimuli; *estimation uncertainty* arising from their imperfect knowledge of the probabilistic relationships governing stimulus transition contingencies within contexts; and *volatility uncertainty* maintained by the unsignalled contextual instability.

The pseudorandom order of TMs was used to generate one stimulus sequence that was used for all participants to ensure comparable learning processes and model parameter estimates. Rest periods occurred every 185 trials, orthogonal to TM switches. The importance of fast responses was stressed. Participants were told that by paying attention to any patterns in the order in which stimuli were presented, and to any switches in these patterns, it may be possible to respond faster. No further information about the nature of the experiment was provided.

Combining the behavioural paradigm with three pharmacological manipulations permitted direct assessment of any separable roles for NA, ACh and DA in belief updating under irreducible uncertainty, estimation uncertainty and volatility uncertainty, and in sensitising the motor system to participants' individual perceptual beliefs. At the end of the experimental session, participants were debriefed, indicated whether they thought they had taken an active drug or placebo, and reported the quality and quantity of their sleep on the previous night (Ellis et al., 1980).

4.3.4 Model-agnostic analyses

Trial-wise RT was calculated as the time between stimulus onset and the subsequent button press. The RT data were log-transformed (Bestmann et al., 2014). A series of conventional, model-agnostic analyses of behaviour were first conducted to assess whether participants learned about the underlying stimulus transition contingencies, and whether learning was influenced by the pharmacological interventions. To assess the interaction between stimulus transition probability and drug, trials were binned according

to three probability levels corresponding to the presented stimuli's true transition probabilities as existed in the TMs (High: 0.85 and 0.70; Mid: 0.25 and 0.20; Low: 0.05) (Galea et al., 2012; Bestmann et al., 2014). A repeated-measures analysis of variance (RM-ANOVA) was used to compare mean log(RTs) for correct responses across these three probability levels and between drug-groups.

To obtain a model-agnostic indication of learning across the course of the probabilistic contexts, a median split was performed on each 50-trial contextual block. A RM-ANOVA was used to compare mean log(RTs) on correct Early (1-25) and Late (26-50) trials at each probability level, and between drug-groups. To assess any learning in more detail, RTs on correct, high probability Early trials were examined after having been baseline-corrected by subtracting the mean RT on the last three correct, high probability trials of the previous context.

In many behavioural response time tasks, participants typically demonstrate post-error slowing, i.e., slower responses on trials following those on which they made an error (Rabbitt, 1966; Botvinick et al., 2001; Gehring and Fencsik, 2001; Cavanagh et al., 2014). To identify any evidence of post-error slowing during the PSRTT, a RM-ANOVA was used to compare log(RTs) on correct trials that immediately followed both correct and erroneous responses. A further RM-ANOVA compared log(RTs) on correct, post-infrequent trials (i.e., trials following those with a true transition probability of 0.05) and correct trials following trials with a true transition probability >0.05.

4.3.5 Model-based analyses

While model-agnostic analyses offer a heuristic indication of learning and possible drug effects, a model-based approach permits quantification of participants' (approximate) inferences and subjective expectations about the transitions, which are driven by data-limited observations. A novel instantiation of the HGF model, consisting of a three-level perceptual model and a response model (see Chapter 3 and Figure 4.2), was therefore applied to the data. Thus, it was possible to map each individual's estimated perceptual beliefs about stimulus transitions, transition contingencies and volatility, and the respective *irreducible*, *estimation* and *volatility uncertainty* about these beliefs, onto his/her observed log(RT) responses. The model was implemented using the 'tapas_logrt_linear_whatworld' code contained in the HGF Toolbox (http://www.translationalneuromodeling.org/tapas/).

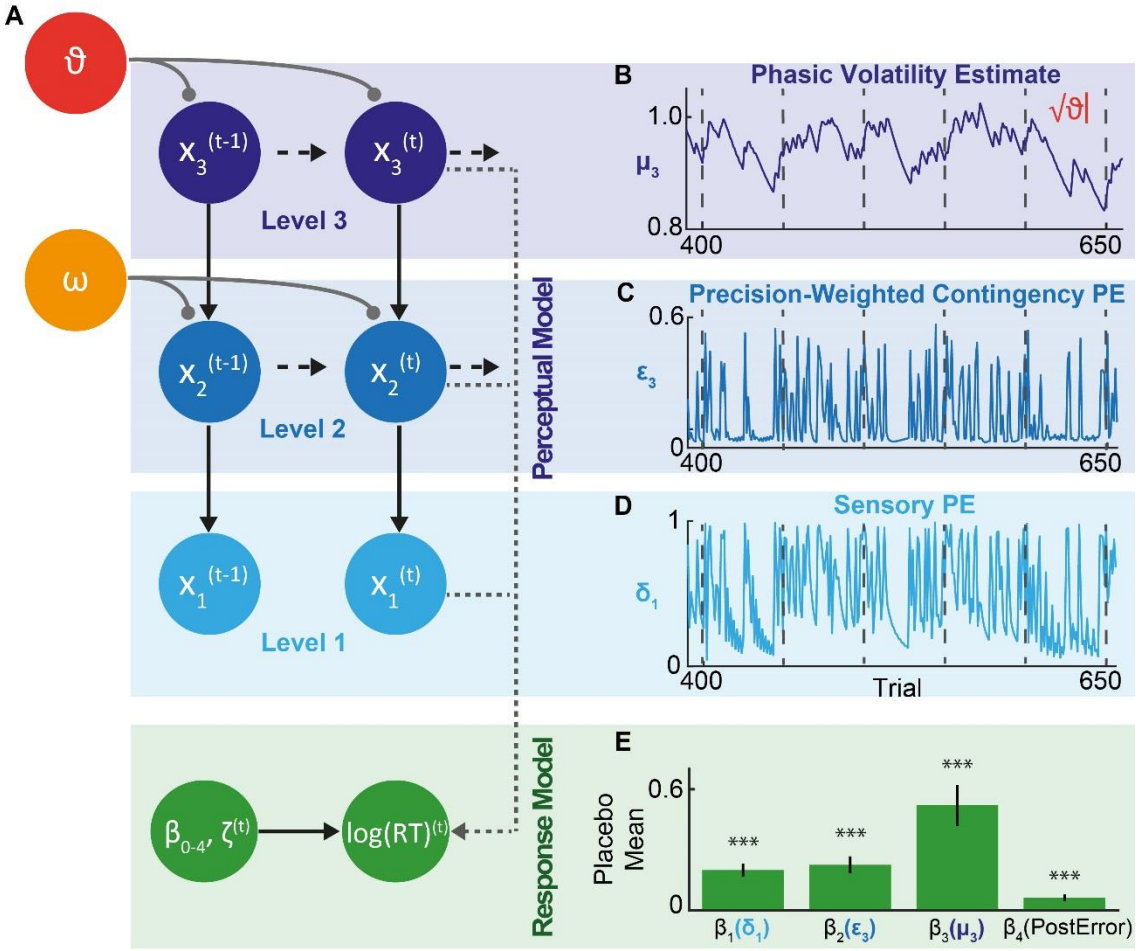

Figure 4.2 The Hierarchical Gaussian Filter (HGF). *(A) The perceptual model tracks an individual's learning of the task's structure across three levels. State* $\boldsymbol{x}_1$ *represents trial-wise transitions from one stimulus to the next,* $\boldsymbol{x}_2$ *the transition contingencies, and* x_3 *the phasic volatility, where t is the current trial number and bold font is used to indicate a matrix. Participants hold and update beliefs about the true quantities at each level, with a mean μ and a variance σ. ϑ and ω are participant-specific parameters that couple the levels and determine the respective speed of belief updating about phasic volatility and transition contingencies. The response model describes the mapping from a participant's trial-wise beliefs onto their observed log(RT) responses. (B) Example of the trial-wise dynamics at level 3 for Placebo Participant 2.* μ_3 *reflects the participant's belief about the true phasic volatility (*x_3*). Vertical dashed lines indicate true context switches.* μ_3 *tends to increase following a context change and then decreases over the course of a context as the participant learns the new contextual rule and thus perceives the environment to be increasingly stable.* $\sqrt{\vartheta}$ *is a variance determining the step-size of* μ_3 *and therefore how quickly the participant updates their phasic volatility estimates. (C) As in B, but for precision-weighted contingency PE (*$\boldsymbol{\varepsilon}_3$*) at level 2. This estimate results from weighting the contingency PE (*$\boldsymbol{\delta}_2$*) by a precision ratio that captures uncertainty about input from*

*the level below relative to the level above. The higher the precision at level 2, the more meaningful a deviation from the predicted stimulus transition contingency. This in turn increases the impact on phasic volatility belief updating at level 3. For simplicity, the depicted ε_3 trajectory is for true transition changes only. (D) As in B and C, but for sensory PE ($\boldsymbol{\delta_1}$) at level 1. This estimate arises from irreducible uncertainty about stimulus transitions. Trial-wise values are equivalent to 1-$\hat{x}_1$, where $\hat{x}_1$ is equal to the probability of the predicted transition. Again, for simplicity only δ_1 values for truly occurring transitions are shown here. (E) Mean β values for the Placebo group indicate that increases in sensory PE (β_1), precision-weighted contingency PE (β_2), and phasic volatility estimates (β_3) slowed participants' trial-wise log(RTs). There was also evidence of post-error slowing (β_4). Results are mean ± SEM. *** p<0.001.*

4.3.5.1 Perceptual model

The perceptual model tracks an individual's learning of the task's structure: the trial-wise stimulus transitions at level 1, the probability of the transitions (i.e., transition contingencies) at level 2, and the volatility of transition contingencies at level 3 (Figure 4.2A). It is hierarchical in that learning not only occurs simultaneously at multiple levels, but belief updating at one level is constrained by beliefs at the level above. This provides a generic framework for implementing dynamic learning rates, which are crucial for learning in volatile environments (Behrens et al., 2007; den Ouden et al., 2010).

Trial-wise trajectories of a participant's perceptual estimates at each level evolve according to the predictions made and outcomes experienced by that individual (Figure 4.2B-D). At levels 2 and 3, these estimates are modelled by Gaussian distributions with a mean (μ) and a variance (σ), the latter reflecting the uncertainty of the estimate. Precision (π) of the estimate is equal to inverse variance (1/σ). *Irreducible uncertainty* at level 1 gives rise to sensory PE, $\boldsymbol{\delta_1}$. *Estimation uncertainty* at level 2 gives rise to contingency PE, $\boldsymbol{\delta_2}$. PEs can be weighted according to their precision (inverse uncertainty). At level 1, this gives us precision-weighted sensory PE, $\boldsymbol{\varepsilon_2}$, and at level 2 precision-weighted contingency PE, $\boldsymbol{\varepsilon_3}$. *Volatility uncertainty* arises from phasic volatility beliefs, μ_3, at level 3. Please refer to Chapter 3 for full details.

Importantly, the HGF does not assume fixed learning across the population but rather contains participant-specific parameters that couple the hierarchical levels and allow for individual expression of approximate Bayes-optimal learning. ϑ determines the speed of learning about volatility, i.e., the rate at which estimates of *phasic volatility (μ_3)* are updated. As such, ϑ encapsulates metavolatility, i.e., the rate at which volatility changes, with higher values implying a belief in a more unstable world and leading to a more

variable learning rate (as expressed in phasic volatility belief updating). By contrast, ω is a constant component of the volatility and captures how rapidly individuals generally update their beliefs about transition contingencies at level 2. Changes in ω therefore lead to a *tonic* alteration of the learning rate. By comparing ϑ and ω estimates for each of the drug-groups to the Placebo group, it was possible to interrogate the effects of NA, ACh and DA antagonism on perceptual belief updating.

4.3.5.2 Response model

The response model describes the mapping from a participant's trial-wise beliefs, as provided by the perceptual model, onto his/her observed responses, log(RTs). I reasoned that there are several variables that could influence trial-wise log(RT). Therefore, three response models were constructed and compared using random effects Bayesian model selection (Stephan et al., 2009; Rigoux et al., 2014) and associated techniques for assessing differences in model frequencies across groups as implemented in the VBA toolbox (Daunizeau et al., 2014). Random effects Bayesian model selection allows for heterogeneity in the population; the best model for each individual is allowed to vary, producing an estimate of model frequency in the population (i.e., for how many participants that model is the best model), and an exceedance probability that the model is the most frequently utilised in the population. This is a more conservative approach than conventional fixed-effects analyses, which assume that data from all participants are best explained by a single model.

Each response model proposed that log(RT) on any given trial is a linear function of a constant component of log(RT) and several other factors. Since there is evidence, both from earlier work (Rabbitt, 1966; Botvinick et al., 2001; Gehring and Fencsik, 2001; Cavanagh et al., 2014) and the present study, that participants' RTs increase on a trial following an incorrect response, post-error slowing was included in each response model. While the perceptual model assumed that participants updated their beliefs according to the stimulus presented on each trial, the response model incorporated correct trials only.

The extra factors in the different models came from quantities at each level of the HGF that might influence log(RT). The first response model contained the following parameters: δ_1 (sensory PE), due to evidence that DA sensitises motor responses to low-level PE (Galea et al., 2012; Bestmann et al., 2014), ε_3 (precision-weighted contingency PE), which has been shown to correlate with activity in the cholinergic basal forebrain (Diaconescu et al., 2017), and μ_3 (estimated phasic volatility), which is relevant to switching tasks for which there is a proposed role for DA (Cools et al., 2001a, 2001b).

For each parameter, the quantity relates to the true stimulus transition on each trial. ζ is Gaussian noise.

Response Model 1:

$$\log(\mathrm{RT})^{(t)} = \beta_0 + \beta_1(\delta_1^{(t)}) + \beta_2(\varepsilon_3^{(t)}) + \beta_3(\mu_3^{(t)}) + \beta_4(\mathrm{PostError}^{(t)}) + \zeta^{(t)}$$

Equation 4.1

Alternative research has indicated that activity in the dopaminergic midbrain correlates with the precision-weighted form of sensory PE, ε_2 (Iglesias et al., 2013; Diaconescu et al., 2017). To disambiguate whether motor responses are modulated according to raw sensory PE or the confidence one has in their sensory predictions, the second response model contained ε_2 instead of δ_1.

Response Model 2:

$$\log(\mathrm{RT})^{(t)} = \beta_0 + \beta_1(\varepsilon_2^{(t)}) + \beta_2(\varepsilon_3^{(t)}) + \beta_3(\mu_3^{(t)}) + \beta_4(\mathrm{PostError}^{(t)}) + \zeta^{(t)}$$

Equation 4.2

Since δ_1 and ε_2 are highly correlated, a third response model containing both parameters was constructed to ascertain whether one had a higher degree of explanatory power in terms of determining log(RT).

Response Model 3:

$$\log(\mathrm{RT})^{(t)} = \beta_0 + \beta_1(\delta_1^{(t)}) + \beta_2(\varepsilon_2^{(t)}) + \beta_3(\varepsilon_3^{(t)}) + \beta_4(\mu_3^{(t)}) + \beta_5(\mathrm{PostError}^{(t)}) + \zeta^{(t)}$$

Equation 4.3

4.3.5.3 Model fitting

For each participant, individual maximum *a posteriori* estimates for perceptual and response model parameters were jointly obtained using the Broyden-Fletcher-Goldfarb-Shanno algorithm as implemented in the HGF Toolbox. Where priors were required, they were defined by inverting the perceptual model in isolation, given the known stimulus sequence (using the function 'tapas_bayes_optimal_whatworld_config'), under suitably uninformative priors. The resulting posterior estimates were then used to define the priors for the subsequent inversion of the full model given the behavioural data (see Table 4.1). In other words, the prior means in the empirical data analysis corresponded to those parameter values for which the stimulus sequence would generate minimal surprise (in an observer with the aforementioned uninformative priors).

Parameter	Notes	Prior	
Perceptual Model			
ϑ	Metavolatility belief parameter; controls the step size of the Gaussian random walk at level 3. Estimated in logit space.	Mean Variance Upper bound	0 2 0.01
ω	Tonic volatility belief parameter; a constant component of the learning rate at level 2.	Mean Variance	-6 25
Stimulus Transitions ($\boldsymbol{x}_1$)	4x4 matrix; the predictions are a sigmoid transformation of the probabilities represented in $\boldsymbol{x}_2$, and so do not have a starting prior value.	$\boldsymbol{\mu}_1$: Mean Variance $\boldsymbol{\sigma}_1$: Mean Variance	 NaN NaN NaN NaN
Stimulus Transition Contingencies ($\boldsymbol{x}_2$)	4x4 matrix; estimated conditional probabilities for the 16 possible stimulus transitions are updated on each trial. At level 2, estimates are made in logit space (-1.0986 is equivalent to a 0.25 probability).	$\boldsymbol{\mu}_2$: Mean Variance $\boldsymbol{\sigma}_2$: Mean Variance	 -1.0986 0 0 log(1)
Volatility (x_3)	Scalar; one trial-wise volatility estimate is updated after each stimulus transition.	μ_3: Mean Variance σ_3: Mean Variance	 1 0.1 log(0.1) 1
Response Model			
β_0	log(RT) constant	Mean Variance	log(500) 3
β_1	Sensory PE (δ_1)	Mean Variance	0 4
β_2	Precision-weighted contingency PE (ε_3)	Mean Variance	0 4
β_3	Volatility estimate (μ_3)	Mean Variance	0 4
β_4	Post-error	Mean Variance	0 3
ζ	Noise	Mean Variance	-3 $1e^{-3}$

Table 4.1 A summary of HGF parameters and priors.

All priors are specified in the space in which they are estimated. For an account of how this relates to the native space of that parameter, please refer to Chapter 3 and to the original description of the model (Mathys et al., 2011).

4.3.5.4 Parameters of interest

Since the HGF separates the relatively complex and interacting factors that influence RTs in a computationally limpid way, the individual effects of the three pharmacological manipulations on perceptual belief updating and response modulation could be probed. By comparing participant-specific perceptual parameters (ϑ and ω) in each drug-group to Placebo, it was possible to characterise learning under pharmacological NA, ACh and DA manipulations. Moreover, any neuromodulatory effects on perceptual belief updating could be distinguished from those on the sensitivity of motor responses (as reflected by the response model β estimates) to perceptual beliefs.

4.3.6 Statistical analyses

In reporting statistical differences, a significance threshold of α=0.05 was used. Where assumptions of sphericity were violated (Mauchly's test $p<0.05$), the Greenhouse-Geisser correction was applied. Since a significant time x drug interaction on self-reported alertness was identified (see section 4.4.5.1 for details), the participant-specific difference in alertness between baseline and the time corresponding to peak drug concentration, *Δalertness*, was used as a covariate in all analyses to control for any inter-participant variability in subjective drug effect.

For comparisons across the four drug-groups, partial eta-squared (η_p^2) is reported as the effect size. The key experimental question pertained how different neuromodulators influence learning and response modulation compared to placebo. Therefore, planned comparisons were made between each of the three active drug-groups (NA-, ACh- and DA-) and the Placebo group by fitting a linear model separately for each participant-specific model parameter (ϑ, ω and each β). Here a Benjamini-Hochberg correction for three pairwise comparisons was applied to account for the false discovery rate (FDR) (Benjamini and Hochberg, 1995). For pairwise comparisons, Cohen's *d* is reported as the effect size.

4.3.7 Control analyses

4.3.7.1 Permutation tests

In addition to the linear models used to assess the effects of the three drug manipulations on the HGF model parameters, permutation tests randomising drug assignment over participants were conducted to make distribution-free comparisons. 10,000 permutations were run per parameter of interest. For each parameter and each permutation, the difference between the mean for each permuted drug and the mean for the permuted

Placebo was calculated. The permutation values were then tested by calculating the fraction of the permutation points with larger absolute differences than, but in the same direction as, those differences observed in the empirical data.

4.3.7.2 Exhaustive response model comparison

To further verify that a response model that offered the best means by which to explain trial-wise log(RT) had been identified, a more exhaustive set of models containing different combinations of parameters from the HGF was compared for the Placebo group. A family-wise model comparison was first run on models containing every combination of the parameters δ_1, ε_2, ε_3 and μ_3 (Family 1) versus models containing every combination of predicted uncertainty estimates from each level of the HGF, $\hat{\sigma}_1$, $\hat{\sigma}_2$ and $\hat{\sigma}$ (Family 2). All models contained an additional parameter for post-error slowing. Once the winning family was identified, random effects Bayesian model selection was run on all models in that family. See section 4.4.5.3 for details.

4.3.7.3 Model parameter correlations

To demonstrate that the HGF provided a good fit to the behavioural data, the correlations between the Bayesian parameter averages (BPAs) for model parameters in each drug-group were assessed.

4.3.7.4 Residuals

For further verification that the HGF model provided a good fit to the behavioural data, the residuals between the observed log(RTs) and those predicted by the model were assessed for each drug-group. To confirm that the HGF did not systematically under- or over-estimate log(RTs) at true contextual change-points, autocorrelations between residuals for participants in each drug-group were calculated.

4.3.7.5 Simulations

To demonstrate that the HGF can capture the effects reported in the results, and to illustrate the implications of different model parameters further, the HGF was used to generate simulated log(RT) data. First, 100 simulations were run for each set of posterior parameter values obtained for each participant in the Placebo group, generating 1200 log(RTs) for each run. The simulated log(RTs) on high (p=0.85 or p=0.70), mid (p=0.25 or p=0.20) and low (p=0.05) probability trials were then averaged, i.e., to mirror the model-agnostic analyses. For each of a series of further simulations, the same parameter settings were taken, but particular parameters of interest, identified based on the empirical observations, were modified. For these parameters of interest, the estimated

parameters for each Placebo participant were shifted by the difference between the Placebo group average and the relevant drug-group average for that parameter. All averaging was performed in the space in which the parameters were estimated. Again, 100 runs for each “computationally drugged” participant were run and the simulated log(RTs) were averaged across three probability levels. Thus, it was possible to assess the impact of different model parameters on log(RT), and to compare simulated log(RTs) to empirical data in each drug-group.

4.4 Results

Data from 124 participants are reported. Four participants were excluded from analyses: three due to high missed response rates (≥11%) and one because behavioural model parameter estimation (using the Broyden-Fletcher-Goldfarb-Shanno algorithm) did not converge. The four drug-groups were matched for gender (Kruskal-Wallis test: H_3=0.53, p=0.912), age (one-way ANOVA: $F_{3,120}$=0.46, p=0.714), body weight ($F_{3,120}$=2.24, p=0.087), education level (H_3=1.31, p=0.727), and all baseline psychometric measures taken (Table 4.2).

	Placebo (n = 32)	NA- (n = 31)	ACh- (n = 29)	DA- (n = 32)	Between-groups difference?
Gender ***(number male)***[#]	13	15	14	14	ns p = 0.912
Age ***(years)***	23.0 ± 4.6	23.1 ± 4.0	22.0 ± 3.6	22.8 ± 4.5	ns p = 0.714
Weight ***(kg)***	61.7 ± 1.5	69.0 ± 2.4	64.9 ± 2.5	64.1 ± 1.7	ns p = 0.087
Education Level ***(1-5)***[#]	2.7 ± 0.2	2.8 ± 0.1	2.6 ± 0.1	2.7 ± 0.1	ns p = 0.727
Digit Span (forwards + ***backwards)***[#]	13.1 ± 0.4	13.0 ± 0.5	12.7 ± 0.4	13.8 ± 0.5	ns p = 0.252
Impulsivity: ***BIS-11***	61.9 ± 1.5	65.4 ± 1.5	64.7 ± 1.6	63.3 ± 2.0	ns p = 0.465
Risk-taking: DOSPERT ***(total)***	104.3 ± 3.0	113.5 ± 3.6	107.7 ± 3.7	105.5 ± 3.5	ns p = 0.245
Distractibility: ***CFQ***	38.3 ± 1.9	41.7 ± 1.6	43.9 ± 1.8	40.3 ± 1.9	ns p = 0.185
Sleep quantity on the previous night ***(hours)***[#]	7.3 ± 0.2	7.2 ± 0.2	6.7 ± 0.3	7.1 ± 0.2	ns p = 0.513
Sleep quality on the previous night ***(1-8)***[#]	5.5 ± 0.3	5.7 ± 0.2	5.3 ± 0.3	5.4 ± 0.2	ns p = 0.620
Fatigue during task ***(0 – 100)***	44.6 ± 3.9	44.8 ± 3.6	43.5 ± 2.6	41.2 ± 3.7	ns p = 0.876
Active drug ***(%)***[#]	50	77	86	44	p = 0.001

Table 4.2 Participant details for each experimental group.

Between-groups comparisons revealed no significant differences (ns = non-significant) for gender, age, body weight, education level, baseline working memory (Digit Span), impulsivity (Barratt Impulsiveness Scale; BIS-11), risk-taking (Domain-Specific Risk-Taking Scale; DOSPERT), distractibility (Cognitive Failures Questionnaire; CFQ), fatigue during the task, or sleep quality or quantity on the previous night. For continuous data, one-way ANOVAs were used to test for any between-group differences. For discrete data (#), Kruskal-Wallis tests were applied. Education Level refers to the highest attained from the following: 1 = compulsory education (≤ 12 years); 2 = further education (13-14 years); 3 = undergraduate degree (15-17 years); 4 = one postgraduate degree (≥ 18 years); 5 = multiple postgraduate degrees. Age data are mean ± SD. Remaining data are mean ± SEM. Active drug refers to the percentage of participants within each group who reported at the end of the experiment that they believed they had received an active drug.

4.4.1 Model-agnostic results

On average, 90.3 ± 0.8% (±SEM), 88.4 ± 1.2%, 87.7 ± 1.3% and 89.2 ± 0.91% of trials were correct in the Placebo, NA-, ACh- and DA- groups respectively. The percentages of correct responses did not differ between groups ($F_{3,123}$=1.12, p=0.345).

First, a RM-ANOVA was conducted on the log(RTs) for correct responses on trials binned according to the five true conditional probabilities that existed in each of the TMs, grouped into High (0.85 and 0.70), Mid (0.25 and 0.20), and Low (0.05) transition probabilities, with drug as a between-subjects factor (Figure 4.3A). This revealed a significant decrease in log(RTs) with increasing transition probability (main effect of probability: $F_{1.27,151.36}$=483.50, p<0.001, effect size η_p^2=0.80), which was modulated by drug-type (probability x drug interaction: $F_{3.82,151.36}$=12.37, p<0.001, η_p^2=0.24), but not by Δalertness (p=0.909).

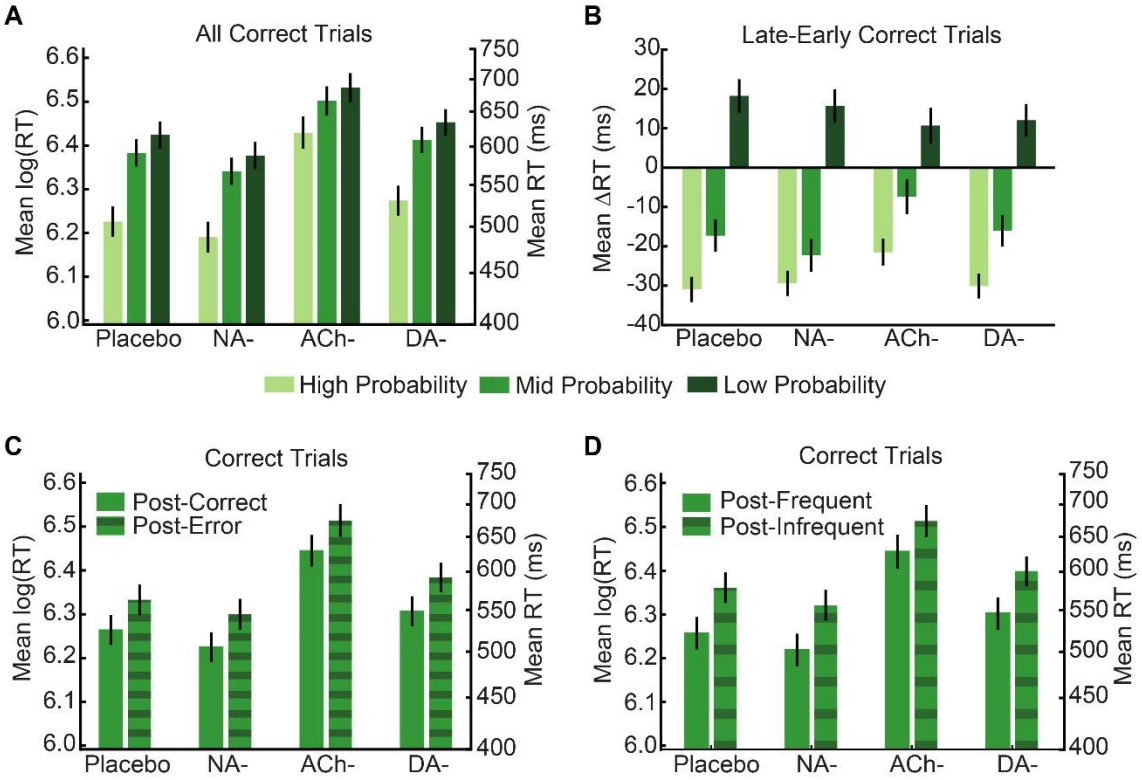

Figure 4.3 Model-agnostic results. *Changes in log(RT) indicate that participants learned to predict the stimulus transitions. (A) In all four groups, log(RT) increased as a stimulus' true transition probability decreased. (B) A median split on each 50-trial contextual block was used to compare mean log(RTs) on Early (1-25) and Late (25-50) trials at each probability level. Over the course of a context, participants became faster at responding to High and Mid probability stimuli, and slower at responding to Low probability stimuli. Raw RTs are plotted here to simplify interpretation of ΔRT, but statistics were run on log(RTs). (C) Across drug-groups, participants showed evidence of post-error slowing on correct trials that followed an erroneous response compared to those following correct responses. (D) Participants also showed evidence of slowing on correct trials that followed an infrequent stimulus transition. Results are mean ± SEM, corrected for Δalertness. Results shown in A, B and D were modulated by drug-group.*

Moreover, across the course of a contextual block (Figure 4.3B), participants became faster at responding to High and Mid probability stimuli and slower at responding to Low probability stimuli (significant main effects of probability: $F_{1.28,152.70}$=476.88, p<0.001, η_p^2=0.80 and time: $F_{1,119}$=12.01, p<0.001, η_p^2=0.34; probability x time interaction: $F_{2,238}$=113.73, p<0.001, η_p^2=0.49). The effect was modulated by drug-type (probability x time x drug interaction: $F_{6,238}$=3.10, p=0.006, η_p^2=0.07), but again not systematically related to differences in Δalertness (all p>0.06). Post-hoc (FDR-corrected) pairwise comparisons indicated that the impact of drug was driven by the ACh- group, which

showed significant log(RT) slowing compared to Placebo (t_{58}=3.06, p=0.009, effect size Cohen's *d*=0.80). Together, these results indicate that participants learned about the true stimulus transition contingencies, and that this learning was modulated by the pharmacological manipulations.

Participants showed evidence of post-error slowing on correct trials following those on which they made an error ($F_{1,119}$=108.25, p<0.001, η_p^2=0.48; Figure 4.3C). This effect was not modulated by drug-group (trial-type x drug interaction: p=0.957), or by Δalertness (p=0.608). Participants also demonstrated significant log(RT) slowing on correct, post-infrequent trials (true transition probability = 0.05) compared to all other correct trials ($F_{1,119}$=441.12, p<0.001, η_p^2=0.79; Figure 4.3D), which was modulated by drug-group ($F_{3,119}$=4.47, p=0.005, η_p^2=0.10) but not by Δalertness (p=0.652). This effect was driven by the ACh- group, with (FDR-corrected) pairwise comparisons revealing significant slowing compared to Placebo (t_{58}=3.44, p=0.003 *d*=0.90). Error rates significantly decreased with increasing transition probability (main effect of probability: $F_{1.54,183.50}$=143.60, p<0.001, η_p^2=0.55). The effect was again modulated by drug-type (probability x drug interaction: $F_{4.63,183.50}$=5.21, p<0.001, η_p^2=0.12), but not by Δalertness (p=0.283). There was no between-subjects effect of drug-group (p=0.776).

4.4.2 Model-based results

4.4.2.1 Perceptual model

Overall, the HGF tracked the true stimulus transitions well (Figure 4.4). Note that the model is uninformed about the true stimulus transition probabilities, but rather bases its estimates on the observed stimulus transitions only.

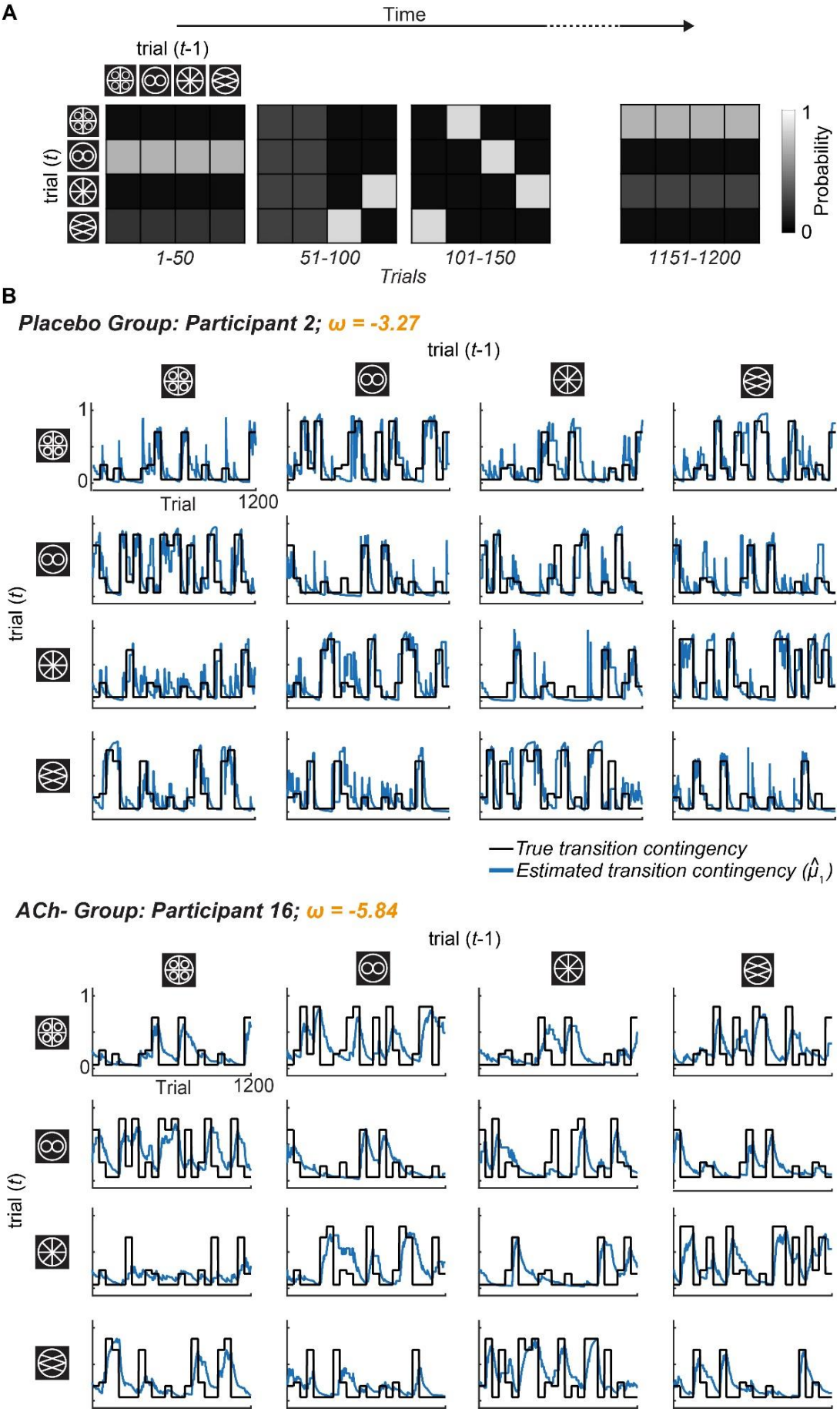
A
Time
trial (t-1)
trial (t)
1
0
Probability
1-50
51-100
101-150
1151-1200
Trials
B
Placebo Group: Participant 2; ω = -3.27
trial (t-1)
1
0
Trial
1200
trial (t)
True transition contingency
Estimated transition contingency ($\hat{\mu}_1$)
ACh- Group: Participant 16; ω = -5.84
trial (t-1)
1
0
Trial
1200
trial (t)

Figure 4.4 Estimated transition contingencies for two example participants. *(A) Transitions between pairs of stimuli, from trial t-1 to trial t, were defined by transition matrices. Every 50 trials the transition matrix switched to a different matrix. (B) Each panel corresponds to one of 16 possible transitions between stimuli across 1200 trials. The black lines indicate the true transition contingencies. The blue lines reflect the participant's inferred estimates (i.e., the posterior expectation of these contingencies, $\hat{\mu}_1$) before seeing the stimulus outcome on each trial. The model tracked the true underlying contingencies and detected change-points. Here, in a representative participant from the Placebo group, the model tracked the true transition contingencies closely, whereas a participant from the ACh- group showed a greater discrepancy in the tracking of the true transition contingencies. This is reflected in the participants' ω estimates: Placebo Participant 2 showed a higher transition contingency learning rate (ω=-3.27) than ACh- Participant 16 (ω=-5.84).*

The punctate change-points contained in the true generative process were detected implicitly by the HGF as an increase in learning rate (α_1; Figure 4.5), which reflects the influence of increased uncertainty and formally corresponds to a reduced contribution of belief precision (denominator in Equation 3.16) to the weighting of PE.

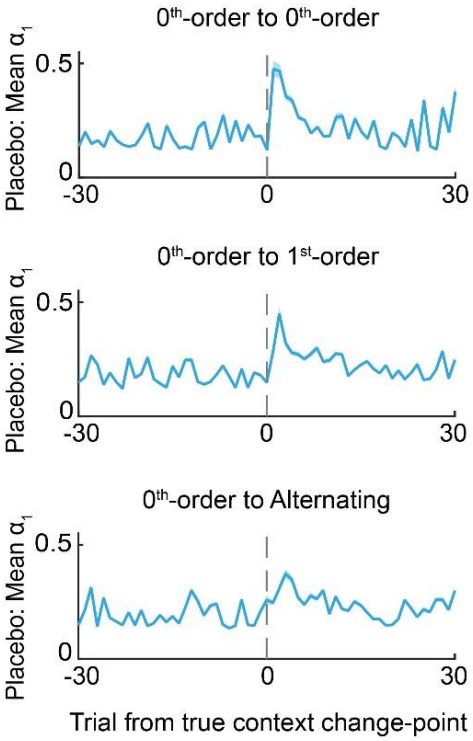

Figure 4.5 Learning rate (α_1) trajectories for the Placebo group. *Increases in α_1 are observed following a true change in context. This α_1 increase is amplified for a more obvious switch from one easy-to-detect 0^{th}-order context to a different 0^{th}-order context.*

In contrast, a switch to an Alternating context, which is trickier to detect, is accompanied by a modest, more gradual increase in α_1. Data are mean ± SEM for truly occurring transitions.

Importantly, when trials were categorised according to participants' trial-wise estimates of transition contingencies, as provided by model parameter $\hat{\mu}_1$ (five bins: 0.8-1, 0.6-0.8, 0.4-0.6, 0.2-0.4, 0-0.2), the same decrease in log(RT) with increasing transition probability found in the model-agnostic results was observed (c.f. Figure 4.6 with Figure 4.3A; significant effect of $\hat{\mu}_1$: $F_{1.68,188.02}$=297.92, p<0.001, η_p^2=0.73). As in the model-agnostic results, this was modulated by drug-group (significant $\hat{\mu}_1$ x drug interaction: $F_{5.04,188.02}$=9.52, p<0.001, η_p^2=0.20), but not by Δalertness (p=0.112).

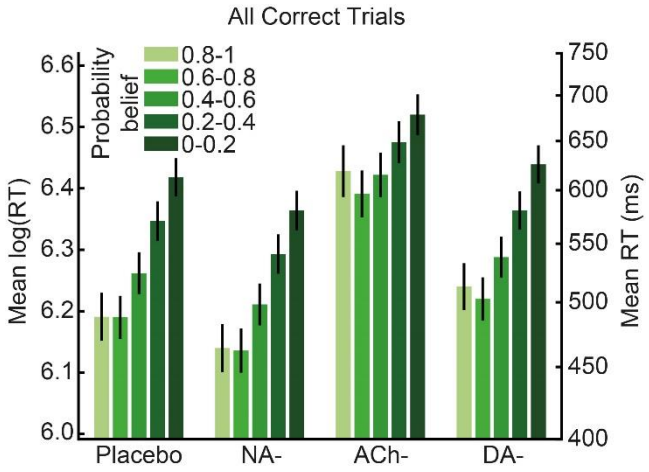

Figure 4.6 Model-based changes in log(RT) mirror the model-agnostic results. *In all four groups, faster responses were observed as participants' estimates of the true transition contingencies increased, demonstrating that the HGF captured the same behavioural effect identified in the model-agnostic analyses, i.e., that participants learned to predict the stimulus transitions and prepared motor responses to high probability transitions (c.f. Figure 4.3A). Results are mean ± SEM, corrected for Δalertness.*

4.4.2.2 Response model

Random effects Bayesian model selection established that Response Model 1 (containing parameters δ_1, ε_3 and μ_3) was superior in all four pharmacological groups by a considerable margin. For the Placebo, NA-, ACh- and DA- groups respectively, the posterior model probabilities were 0.911, 0.828, 0.636 and 0.829; protected exceedance probabilities (i.e., the probability that Response Model 1 is more likely than any other model in the comparison set) were 1.000, 1.000, 0.963, 1.000 (Figure 4.7). Moreover, no significant difference in model frequencies between the Placebo group and any of the drug-groups was identified (NA- vs Placebo: p=0.958, ACh- vs Placebo: p=0.560, DA-

vs Placebo: p=0.955). Therefore, Response Model 1 was used for all subsequent analyses.

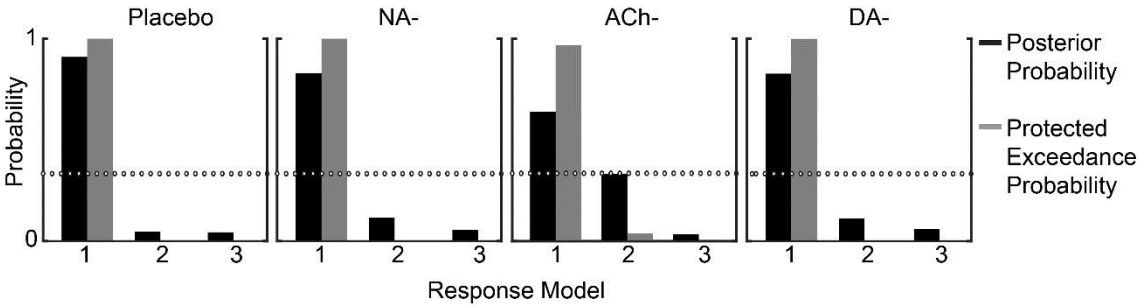

Figure 4.7 Random effects Bayesian model selection results. *Response Model 1 was found to be superior in all four groups. Posterior probabilities quantify the likelihood of each model given the data. Protected exceedance probabilities quantify how likely it is that any given model is more frequently utilised by individuals than all other models in the comparison set, while also protecting against the possibility that the observed variability in (log-) model evidences could be due to chance. The dotted line indicates the threshold for chance-level posterior probabilities (p=0.33).*

All regression coefficients for the Placebo group were significantly greater than 0 (Figure 4.2E), meaning that sensory PE ($\beta_1(\delta_1)$: t_{30}=7.90, p<0.001, effect size *d*=1.41), precision-weighted contingency PE ($\beta_2(\varepsilon_3)$: t_{30}=6.33, p<0.001, *d*=1.13) and phasic volatility estimates ($\beta_3(\mu_3)$: t_{30}=5.49, p<0.001, *d*=0.98) all had slowing influences on log(RT), and that there was evidence of post-error slowing (β_4(PostError): t_{30}=5.85, p<0.001, *d*=1.05). Each of the drug-groups showed equivalent post-error slowing to the Placebo group (all p>0.54; Figure 4.9F), mirroring the model-agnostic result. The lack of a difference in the noise parameter ζ between the Placebo group and any of the drug-groups (all p>0.34; Figure 4.9A) indicates that the model's ability to predict log(RT) was unaltered under the drug manipulations.

4.4.3 The influence of noradrenaline and acetylcholine in perceptual uncertainty computations

4.4.3.1 Noradrenaline antagonism increased phasic volatility learning rate

The noradrenergic (α1-adrenoceptor) antagonist prazosin increased the rate at which individuals updated their volatility estimates, as reflected by an increase in ϑ (linear model: t_{60}=2.32, p=0.033, effect size Cohen's *d*=0.60; Figure 4.8A). A higher ϑ leads to greater fluctuations in participants' phasic volatility estimates, μ_3, resulting in a more variable phasic learning rate. By contrast, there was no effect on ω (p=0.388; Figure

4.8B), indicating that the tonic learning rate about the probabilistic contexts remained unchanged.

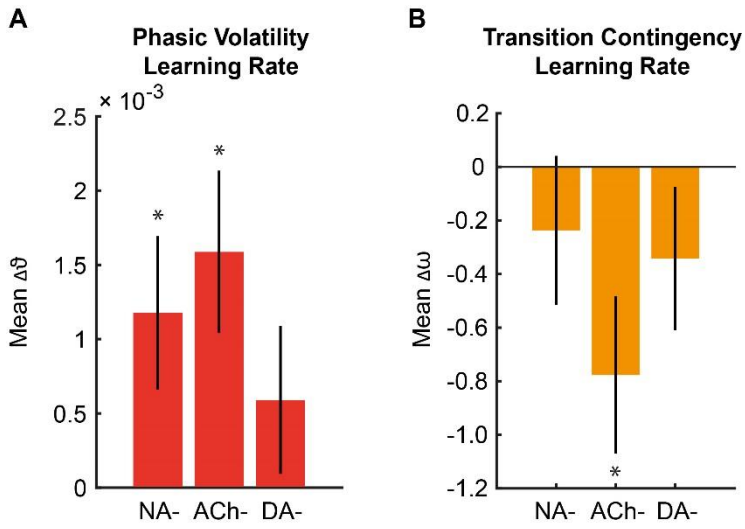

Figure 4.8 Perceptual model parameter results. *(A-B) Compared to the Placebo group, NA and ACh antagonism modulated participants' perceptual belief updating. NA- increased the rate at which participants updated their phasic volatility estimates (increased ϑ). ACh- decreased the rate at which participants learned about stimulus transition contingencies (decreased ω), and increased the rate at which participants updated their phasic volatility estimates (increased ϑ). Results are (mean Drug) – (mean Placebo), ± the standard error of the difference (SED) between the means of the two samples, and corrected for Δalertness. * p<0.05 following an FDR correction for three multiple comparisons. See Table 4.3 for Placebo group means.*

4.4.3.2 Acetylcholine antagonism slowed learning about stimulus transition contingencies

Muscarinic cholinergic (M1-receptor) antagonism under biperiden had more widespread perceptual effects. While ϑ was again significantly increased compared to Placebo (t_{58}=2.95, p=0.012 *d*=0.81; Figure 4.8A), ω estimates in the ACh- group were significantly reduced (t_{58}=-2.68, p=0.025, *d*=-0.74; Figure 4.8B). The lower estimate of ω indicates that participants were slower to update their transition contingency estimates under biperiden and thus slower to adapt to the probabilistic contexts.

4.4.3.3 Dopamine antagonism had no effect on learning about task structure

The D1/D2 dopamine receptor antagonist haloperidol did not influence the rate at which participants learned about the task's volatility or contextual transition contingencies compared to Placebo (ϑ and ω: both p>0.23).

To summarise, both NA and ACh antagonism altered learning of uncertain events arising from unexpected contextual changes in the environment. Only ACh antagonism disrupted learning of transition contingencies within probabilistic contexts.

4.4.4 Neuromodulatory effects on response modulation

4.4.4.1 Noradrenaline antagonism had no influence on responses

The response model output revealed no significant effects of NA antagonism on participants' capacity to modulate their motor responses according to their perceptual estimates of uncertainty (all p>0.09; Figure 4.9C-E).

4.4.4.2 Acetylcholine antagonism reduced response sensitivity to perceptual beliefs

Compared to Placebo, ACh antagonism reduced the sensitivity of participants' motor responses to sensory PE (β_1**:** t_{58}=-3.27, p=0.004, *d*=-0.90), precision-weighted contingency PE (β_2**:** t_{58}=-2.67, p=0.026, *d*=-0.74) and phasic volatility estimates (β_3**:** t_{58}=-3.95, p<0.001, *d*=-1.09) (Figure 4.9C-E).

4.4.4.3 Dopamine antagonism reduced response sensitivity to phasic volatility

Compared to Placebo, DA antagonism led to a decrease in the influence of phasic volatility estimates on log(RT) (β_3**:** t_{61}=-2.69, p=0.012, *d*=-0.67; Figure 4.9E). This indicates that DA antagonism suppressed the sensitivity of motor responses to higher-level inference. There was no significant effect of DA antagonism on the sensitivity of motor responses to sensory PE or precision-weighted contingency PE (all p≥0.14).

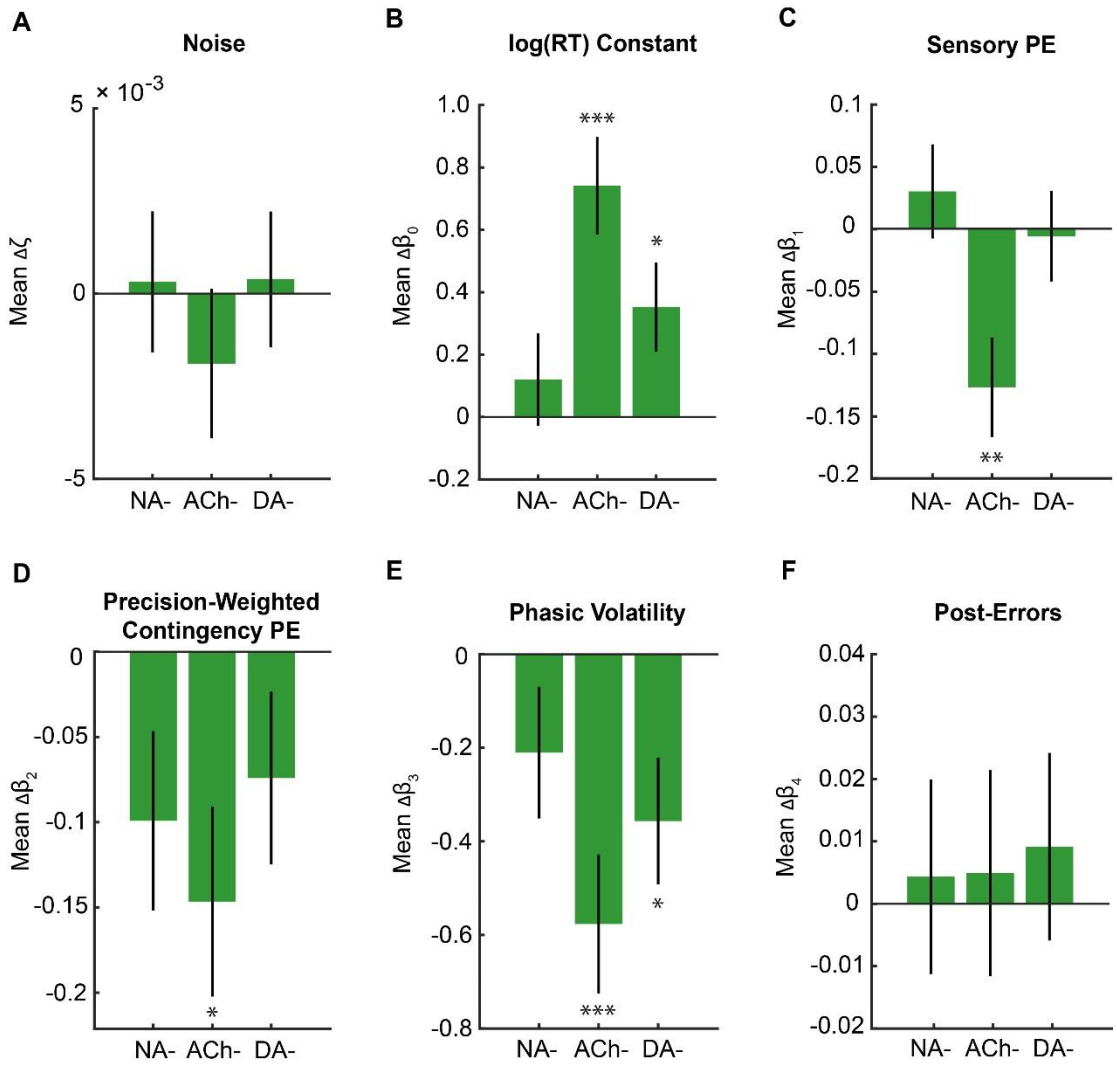

Figure 4.9 Response model results. *DA- antagonism decreased the sensitivity of participants' trial-wise responses to their phasic volatility estimates (β_3). DA- and ACh- antagonism also caused some general response slowing (β_0). The three drug-groups and the Placebo group showed equivalent post-error slowing (β_4) and Gaussian noise (ζ). Results are (mean Drug) – (mean Placebo), ± SED and corrected for Δalertness. * p<0.05, ** p<0.01, *** p<0.001 following an FDR correction for multiple comparisons. See*

Table 4.3 for Placebo group means.

In addition to the effects reported above, the log(RT) constant output indicated that suppressing DA and ACh also led to some general log(RT) slowing (β_0**:** t_{61}=2.54, p=0.019 *d*=0.64; t_{58}=4.85, p<0.001, *d*=1.34 respectively; Figure 4.9B). Subjective Δalertness systematically modulated the effects observed on ϑ (t_{119}=2.54, p=0.013, *d*=0.02), sensory PE (β_1**:** t_{119}=2.53, p=0.013, *d*=0.02) and precision-weighted contingency PE (β_2**:** t_{119}=-3.09, p=0.002, *d*=-0.02).

Parameter	Mean	SEM	t-value	p-value
ϑ	0.0034	0.0035	-	-
ω	-3.5849	0.1880	-	-
β_0	5.7079	0.0992	57.567	<0.001
$\beta_1(\delta_1)$	0.1996	0.0253	7.9020	<0.001
$\beta_2(\varepsilon_3)$	0.2256	0.0357	6.3260	<0.001
$\beta_3(\mu_3)$	0.5208	0.0949	5.4900	<0.001
β_4(Post-Error)	0.0620	0.0106	5.8520	<0.001
ζ	0.0536	0.0013	-	-

Table 4.3 Average perceptual and response model parameters for the Placebo group. *β_0 reflects a constant component of log(RT). $\beta_{1\text{-}4}$ reflect the influence of sensory PE (δ_1), precision-weighted contingency PE (ε_3), phasic volatility estimates (μ_3) and post-error trials on log(RT). All β values were significantly greater than zero (all p<0.001), indicating that these parameters slowed log(RT). All data are corrected for Δalertness.*

4.4.5 Control analyses

4.4.5.1 Physiological and subjective control measures

Self-reported ratings for alertness, calmness and contentedness all changed significantly over the course of the experiment ($F_{1.79,214.39}$=71.60, p<0.001, η_p^2=0.37; $F_{1.88,225.25}$=5.96, p=0.004, η_p^2=0.05 and $F_{2,240}$=25.65, p<0.001, η_p^2=0.18 respectively), but only alertness ratings showed a significant time x drug interaction ($F_{5.36,214.39}$=6.40, p<0.001, η_p^2=0.14). On average, alertness decreased within-participants over the course of the experiment in all four groups. A one-way ANOVA with drug as a between-subjects factor revealed that the degree to which alertness decreased between Baseline (Figure 4.1A) and the time corresponding to peak drug concentration (Post-Drug) varied between groups ($F_{3,120}$=7.92, p<0.001, η_p^2=0.17). More specifically, compared to Placebo, the alertness-decrease was significantly more pronounced in the ACh- and NA- groups (t_{59}=-4.31, p<0.001, *d*=-1.11 and t_{61}=-2.76, p=0.007, *d*=-0.70 respectively).

Heart rate (HR) varied significantly with time ($F_{1.89,\ 226.71}$=129.25, p<0.001, η_p^2=0.52) and this effect was modulated by drug-group ($F_{5.67,226.71}$=5.40, p<0.001, η_p^2=0.12). On average, all groups showed participant-specific HR decreases between Baseline and Post-Drug. The magnitude of HR deceleration differed between groups ($F_{3,120}$=6.65, p<0.001, η_p^2=0.14), but only in the ACh- group was HR deceleration more pronounced than Placebo (t_{59}=-3.14, p=0.002, *d*=-0.81). While systolic blood pressure (BP) varied with time ($F_{2,240}$=7.12, p=0.001, η_p^2=0.06), there was no time x drug interaction

($F_{6,240}$=1.55, p=0.16). Diastolic BP showed no main effect of time ($F_{2,240}$=0.37, p=0.695), but there was a significant time x drug interaction ($F_{6,240}$=3.52, p=0.002, η_p^2=0.08). More precisely, participant-specific differences in diastolic BP between Baseline and Post-Drug varied significantly between groups ($F_{3,120}$=5.11, p=0.002, η_p^2=0.11) due to a significant decrease in diastolic BP in the NA- group compared to the Placebo group (t_{61}=-3.49, p<0.001, *d*=-0.88). This is unsurprising given that the NA- drug administered (prazosin) is used clinically as an anti-hypertensive. A summary of the subjective and physiological measures is reported in Table 4.4.

		Placebo	NA-	ACh-	DA-
Alertness	Baseline	64.4 ± 2.8	64.7 ± 2.5	65.2 ± 2.5	65.1 ± 2.3
	Post-Drug	60.7 ± 2.3	52.5 ± 2.6	48.0 ± 2.3	59.6 ± 3.0
	Post-Task	59.5 ± 2.8	48.7 ± 2.8	49.9 ± 2.4	52.8 ± 3.6
Calmness	Baseline	66.9 ± 2.9	68.9 ± 3.2	66.4 ± 2.6	61.5 ± 3.0
	Post-Drug	70.6 ± 2.6	68.5 ± 2.9	62.9 ± 2.6	67.3 ± 2.4
	Post-Task	60.8 ± 2.9	66.4 ± 2.8	60.2 ± 2.7	63.5 ± 2.5
Contentedness	Baseline	71.2 ± 2.3	70.2 ± 2.6	69.4 ± 2.1	69.5 ± 2.3
	Post-Drug	69.7 ± 2.2	67.2 ± 2.5	63.5 ± 2.1	67.7 ± 2.3
	Post-Task	65.9 ± 2.4	66.7 ± 2.4	59.5 ± 2.1	64.9 ± 2.2
HR	Baseline	69.8 ± 1.9	78.7 ± 2.4	71.0 ± 1.8	74.8 ± 2.1
	Post-Drug	61.2 ± 1.7	72.7 ± 2.3	55.8 ± 1.7	65.3 ± 1.7
	Post-Task	62.8 ± 1.6	73.1 ± 2.0	56.6 ± 1.5	66.7 ± 1.8
Systolic BP	Baseline	110.8 ± 1.9	121.7 ± 2.2	111.5 ± 2.0	117.6 ± 2.6
	Post-Drug	109.5 ± 1.5	117.5 ± 2.1	109.8 ± 2.1	114.5 ± 2.0
	Post-Task	110.1 ± 1.5	121.6 ± 2.1	116.3 ± 2.8	116.4 ± 2.2
Diastolic BP	Baseline	68.8 ± 1.3	73.4 ± 1.1	69.1 ± 1.4	69.6 ± 1.7
	Post-Drug	70.5 ± 1.4	69.0 ± 1.5	68.7 ± 1.5	70.9 ± 1.4
	Post-Task	71.7 ± 1.4	69.7 ± 1.3	70.6 ± 1.9	69.3 ± 1.8

Table 4.4 Subjective and physiological measures for each experimental group. *Readings were taken at baseline, immediately before participants started the PSRTT (i.e., when the drugs were at their most active; Post-Drug), and after completing the PSRTT (Post-Task). Data are mean ± SEM.*

4.4.5.2 Permutation tests

Aside from the effect of ACh- on β_2, all significant effects observed in the multiple comparisons reported above (c.f. Figure 4.8 and Figure 4.9) were mirrored in the results of the permutation tests.

	Comparison (drug vs Placebo)	**Direction of effect (for drug vs Placebo)**	**p-value**	**Significant?**
ϑ	NA- vs Placebo	↑	0.040	*
	ACh- vs Placebo	↑	0.016	*
	DA- vs Placebo	-	0.147	ns
ω	NA- vs Placebo	-	0.203	ns
	ACh- vs Placebo	↓	0.002	**
	DA- vs Placebo	-	00106	ns
β_0	NA- vs Placebo	-	0.322	ns
	ACh- vs Placebo	↑	<0.001	***
	DA- vs Placebo	↑	0.011	*
β_1	NA- vs Placebo	-	0.568	ns
	ACh- vs Placebo	↓	<0.001	***
	DA- vs Placebo	-	0.392	ns
β_2	NA- vs Placebo	-	0.130	ns
	ACh- vs Placebo	-	0.057	ns
	DA- vs Placebo	-	0.104	ns
β_3	NA- vs Placebo	-	0.122	ns
	ACh- vs Placebo	↓	<0.001	***
	DA- vs Placebo	↓	0.006	**
β_4	NA- vs Placebo	-	0.554	ns
	ACh- vs Placebo	-	0.542	ns
	DA- vs Placebo	-	0.711	ns
ζ	NA- vs Placebo	-	0.571	ns
	ACh- vs Placebo	-	0.098	ns
	DA- vs Placebo	-	0.505	ns

Table 4.5 Permutation test results. *10,000 permutations were run for each of the HGF model parameters, randomising drug assignment over participants. Aside from the effect of ACh- on β_2, all significant effects observed in the multiple comparisons reported above*

*(c.f. Figure 4.8 and Figure 4.9) were mirrored in the results of the permutation tests. * p<0.05, ** p<0.01, *** p<0.001; ns = non-significant.*

4.4.5.3 Exhaustive response model comparison

A family-wise model comparison established that Family 1 (containing models with different combinations of the parameters δ_1, ε_2, ε_3 and μ_3) was superior to Family 2 (containing models with different combinations of the parameters $\hat{\sigma}_1$, $\hat{\sigma}_2$ and $\hat{\sigma}_3$) (posterior probability: 0.700; exceedance probability: 0.999; Table 4.6). Mirroring the model comparison results reported above (section 4.4.2.2), random effects Bayesian model comparison on all models in Family 1 identified Response Model 1 as superior (posterior probability: 0.270, protected exceedance probability: 0.844; Table 4.7).

	Model Parameters	**Posterior Probability**	**Exceedance Probability**
Family 1	δ_1, ε_3, μ_3, *PostError*, ζ ε_2, ε_3, μ_3, *PostError*, ζ δ_1, ε_2, ε_3, μ_3, *PostError*, ζ δ_1, *PostError*, ζ δ_1, ε_2, *PostError*, ζ δ_1, ε_2, ε_3, *PostError*, ζ δ_1, ε_2, μ_3, *PostError*, ζ δ_1, ε_3, *PostError*, ζ δ_1, μ_3, *PostError*, ζ ε_2, *PostError*, ζ ε_2, ε_3, *PostError*, ζ ε_2, μ_3, *PostError*, ζ ε_3, *PostError*, ζ ε_3, μ_3, *PostError*, ζ μ_3, *PostError*, ζ	0.700	0.999
Family 2	$\hat{\sigma}_1$, *PostError*, ζ $\hat{\sigma}_2$, *PostError*, ζ $\hat{\sigma}_3$, *PostError*, ζ $\hat{\sigma}_1, \hat{\sigma}_2$, *PostError*, ζ $\hat{\sigma}_2, \hat{\sigma}_3$, *PostError*, ζ $\hat{\sigma}_1, \hat{\sigma}_2$, $\hat{\sigma}_3$, *PostError*, ζ	0.300	0.001

Table 4.6 Results of family-wise Bayesian model comparison. *To further verify that*

Response Model 1 offered the best means by which to explain trial-wise log(RT), a more exhaustive set of linear response models containing different combinations of parameters from the HGF were compared for the Placebo group. A family-wise model comparison was first run on models containing every combination of the parameters δ_1, ε_2, ε_3 and μ_3 (Family 1) versus models containing every combination of $\hat{\sigma}_1$, $\hat{\sigma}_2$ and $\hat{\sigma}_3$ (Family 2). Note that the quantities corresponded to the true transition that occurred on each trial. All models included post-error slowing. Family 1 was found to be superior (posterior probability: 0.700; exceedance probability: 0.999).

Model Number	Model Parameters	Posterior Probability	Protected Exceedance Probability
1	δ_1, ε_3, μ_3, *PostError*, ζ	0.270	0.844
2	ε_2, ε_3, μ_3, *PostError*, ζ	0.023	0.000
3	δ_1, ε_2, ε_3, μ_3, *PostError*, ζ	0.024	0.000
4	δ_1, *PostError*, ζ	0.115	0.026
5	δ_1, ε_2, *PostError*, ζ	0.023	0.000
6	δ_1, ε_2, ε_3, *PostError*, ζ	0.022	0.000
7	δ_1, ε_2, μ_3, *PostError*, ζ	0.022	0.000
8	δ_1, ε_3, *PostError*, ζ	0.024	0.000
9	δ_1, μ_3, *PostError*, ζ	0.144	0.068
10	ε_2, *PostError*, ζ	0.023	0.000
11	ε_2, ε_3, *PostError*, ζ	0.038	0.000
12	ε_2, μ_3, *PostError*, ζ	0.021	0.000
13	ε_3, *PostError*, ζ	0.097	0.013
14	ε_3, μ_3, *PostError*, ζ	0.133	0.049
15	μ_3, *PostError*, ζ	0.021	0.000

Table 4.7 Bayesian model comparison results for Family 1. *Each model contained a combination of the parameters δ_1, ε_2, ε_3 and μ_3, and a parameter for post-error slowing. Response Model 1 was again found to be superior (posterior probability: 0.270; protected exceedance probability: 0.844).*

4.4.5.4 Model parameter correlations

Aside from two exceptions, Bayesian parameter averages (BPAs) for the different model parameters were only moderately correlated across groups (all absolute r≤0.660; Figure

4.10). Higher correlations existed between BPAs for ω (transition contingency learning rate) and μ_3_0 (the initial phasic volatility estimate) (r=-0.948, -0.764, -0.771, -0.983 for Placebo, NA-, ACh- and DA- respectively). This is to be expected on theoretical grounds because the two parameters perform very similar functions in the generative model. Note that the initial value of μ_3 was estimated as μ_3 was used as a predictor of log(RT) in the response model. However, when μ_3_0 was fixed there were no changes to any of the reported main effects.

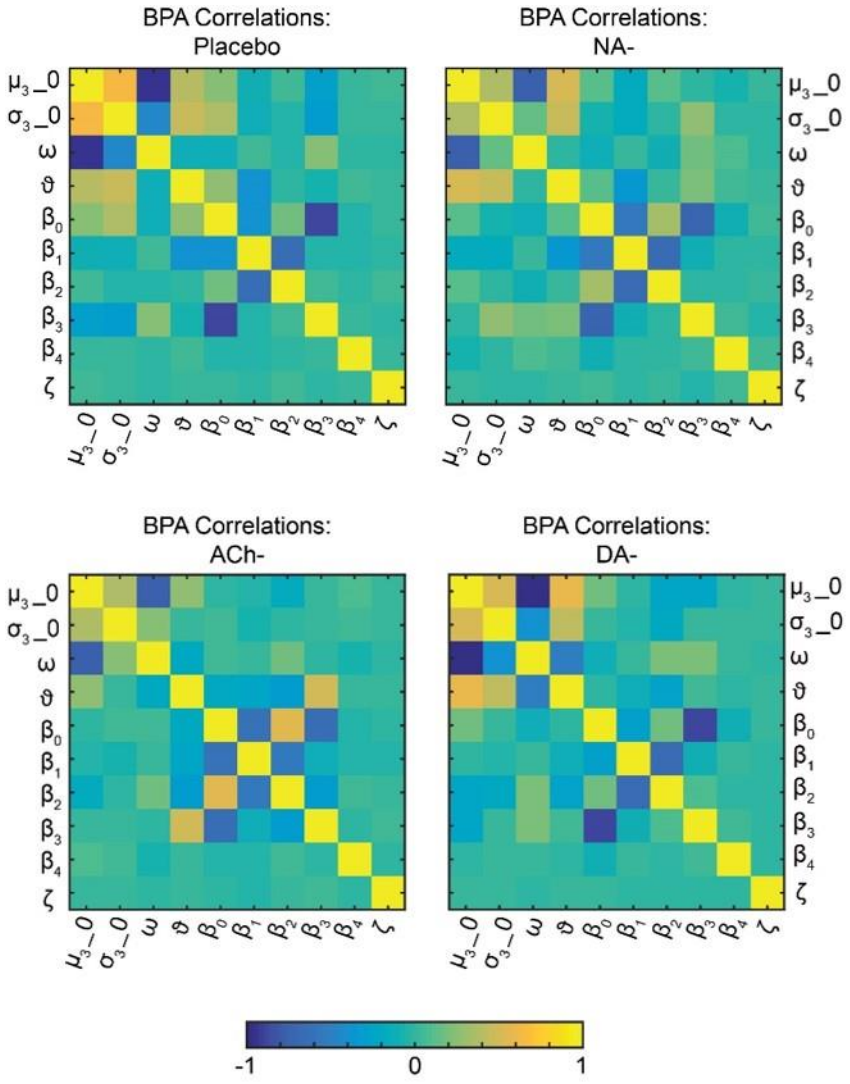

Figure 4.10 Model parameter correlations for the Bayesian parameter averages (BPAs). *Note that μ_3_0 and σ_3_0 are the initial values of μ_3 (the phasic volatility estimate) and σ_3 (the uncertainty about the phasic volatility estimate) respectively.*

Higher correlations also occurred between BPAs for β_0 (log(RT) constant) and $\beta_3(\mu_3)$ (the sensitivity of log(RTs) to phasic volatility estimates) (r=-0.877, -0.736, -0.630 and -0.880). Here the negative correlation indicates that both the constant component of log(RT) and

phasic volatility estimates had a similar slowing effect on log(RT). This reflects the fact that, while including μ_3 as a predictor of log(RT) significantly improves model evidence, it is much less variable than the other predictors because volatility inevitably changes at a slower time scale than transition contingencies.

4.4.5.5 Residuals

The distribution of residuals between the observed log(RTs) and those predicted by the HGF suggests that the model captured the patterns in the data well, and thus provided a good fit to the behavioural data (Figure 4.11). The mean (± SEM) correlations between observed log(RTs) and predicted log(RTs) were 0.38 ± 0.02, 0.36 ± 0.02, 0.26 ± 0.01 and 0.36 ± 0.02 for Placebo, NA-, ACh- and DA- respectively.

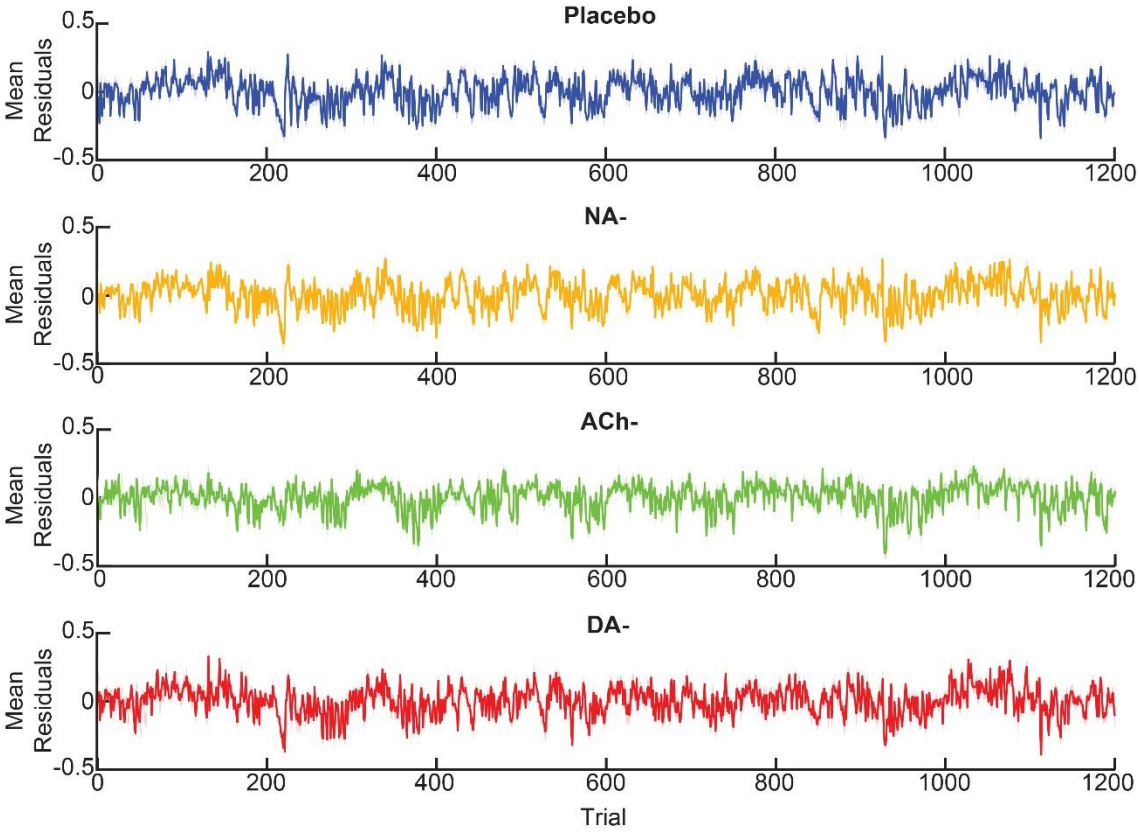

Figure 4.11 Residuals between observed and predicted log(RTs). *The distribution of residuals between observed log(RTs) and those predicted by the HGF suggests that, across drug-groups, the model captured any patterns in the data well. Data are mean ± SEM.*

Moreover, autocorrelations between residuals for participants in each drug-group indicate that the model did not systematically under- or over-estimate log(RTs) at true contextual change-points (Figure 4.12).

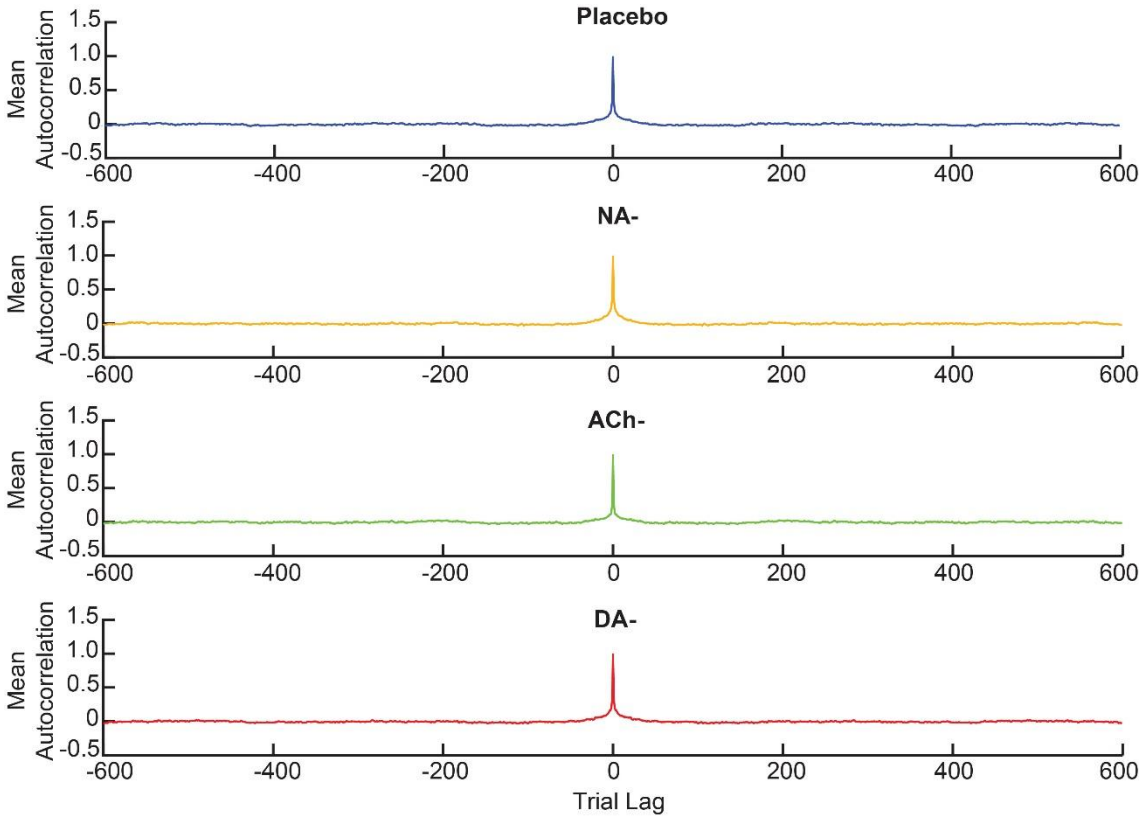

Figure 4.12 Autocorrelations between residuals across trials. *Across groups, the HGF did not systematically under- or overestimate log(RTs) at true change-points. Data are mean ± SEM.*

4.4.5.6 Simulations

Simulated log(RT) data generated using the posteriors for each participant in the Placebo group as model parameters faithfully reflected the increase in log(RT) with decreasing stimulus transition probability that was observed in the Placebo group's empirical data (Figure 4.13). Shifting the parameters significantly altered by the different drug manipulations by the difference between the Placebo group mean for those parameters and the relevant drug-group mean simulated log(RT) data comparable to the empirical data observed in each drug-group. Indeed, simulating NA antagonism by increasing ϑ generated log(RTs) comparable to those for the NA- group. The same was true when DA antagonism was simulated by simultaneously increasing β_0 and decreasing β_3. Similarly, simulating ACh antagonism by increasing ϑ and β_0, and decreasing ω, β_1, β_2 and β_3 produced slower simulated log(RTs) that faithfully reflected the empirical ACh-log(RT) data. Note that, unlike the empirical data, there is no additional slowing caused by post-error effects in the simulated data.

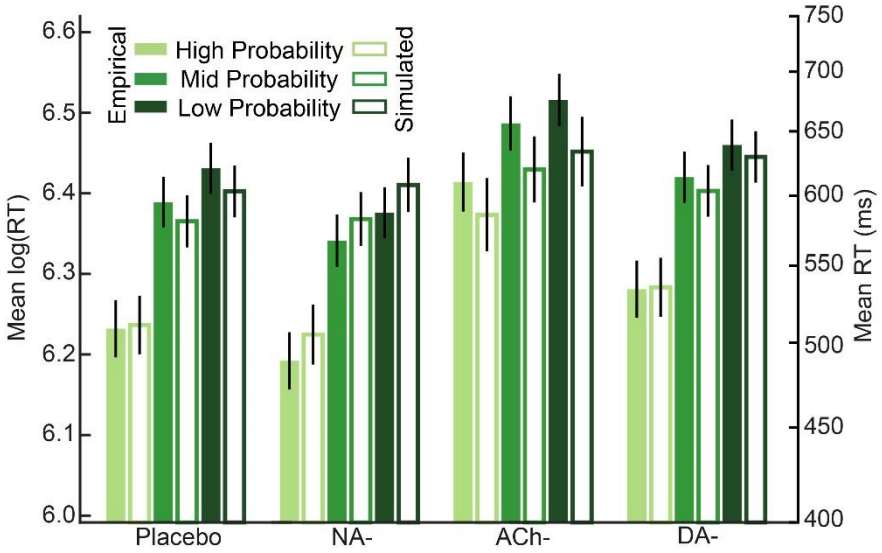

Figure 4.13 Empirical and simulated log(RTs). *Empirical data (filled bars) indicated that log(RT) increased as a stimulus' true transition probability decreased. Simulated data (unfilled bars) generated for the Placebo group and for three "computationally drugged" groups faithfully reflected the empirical log(RTs) in each drug-group. Note that there are no post-error slowing effects in the simulated data. Data are mean ± SEM.*

The simulated data was also able to capture the increase in RT observed in the empirical data following true contextual change-points, as well as the learning that occurred across trials within a stable context. (Figure 4.14). The model-agnostic data shown in Figure 4.14A is for mean (±SEM) ΔRT (collapsed across TM-types) for high-probability trials ($p>=0.70$, as defined in the relevant TM) on which participants made a correct response, following true change-points for each of the drug-groups. The trial-wise ΔRT measure is the difference between RT on each post-change trial and the average of the last three high probability, correct trials in the previous context. RTs increased on the trial following a true change-point across drug-groups (one-way ANOVA with Δalertness covariate: $F_{4,119}=6.52$, $p<0.001$, $\eta_p^2=0.18$), with an additional between-subjects effect of drug-group ($F_{4,119}=7.50$, $p<0.001$, $\eta_p^2=0.16$). Post-hoc comparisons (FDR-corrected) demonstrated that this RT increase was significantly attenuated in the ACh- group compared to the Placebo ($t_{58}=-4.14$, $p<0.001$, $d=-1.09$) group. This is in line with the earlier assessment that individuals in the ACh- group showed poorer learning of the contextual transition contingencies.

Moreover, over the course of the context, there was a decrease in RT for high-probability trials, reflecting learning of the new context. Applying an ANOVA, with a Δalertness covariate, to compare the lines of best fit for each participant's ΔRTs, demonstrated learning across the course of the contexts (reflected by the negative slopes; effect of

slope: $F_{4,119}$=6.52, p<0.001, η_p^2=0.18) and modulation by drug-group (effect of drug: $F_{3,118}$=6.31, p=0.001, η_p^2=0.14). Again, corrected post-hoc comparisons indicated slower learning in the ACh- group compared to the Placebo group (t_{58}=4.05, p<0.001, *d*=1.06) group, in line with the finding of a reduced transition contingency learning rate (as reflected by model parameter ω) following ACh antagonism.

Figure 4.14B indicates that the simulated data echoes the model-agnostic results: there are equivalent between-group differences, most notably a dampened RT increase following a true contextual switch, as well as a reduced learning rate, for the ACh- group.

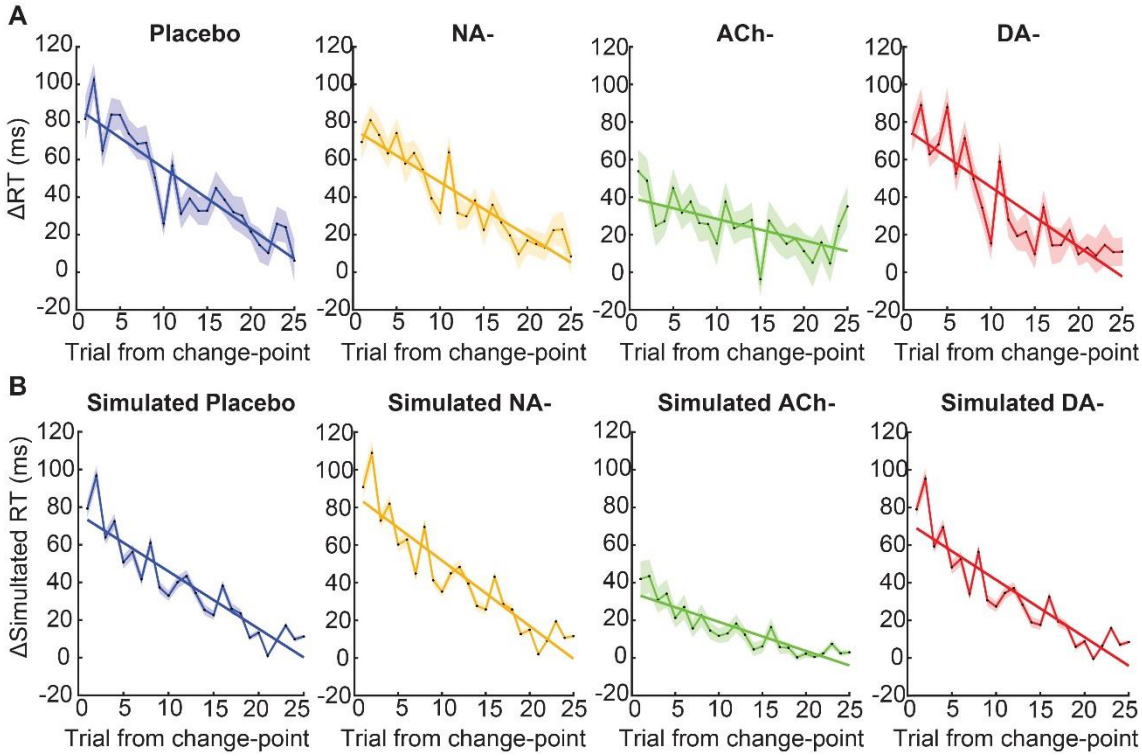

Figure 4.14 Empirical and simulated responses following true change-points. *(A) Mean baseline-corrected RTs for the first 25 (high probability) trials in each context, where the baseline is the mean RT of the last three high probability trials in the previous context. RTs increase following a true contextual change-point, but fall as participants learn the new contextual rule. (B) As in A, but for simulated RTs. The model neatly captures the increase in RTs following true change-points, the reduction in RT that occurs with learning across the course of the new context, and the suppressed effects of change-points and learning on RTs in the ACh- group. Data are mean ± SEM. Raw RTs have been used here to simplify interpretation of ΔRT.*

4.5 Discussion

By implementing a novel PSRTT in conjunction with three pharmacological manipulations and placebo, it was possible to characterise the roles of three neuromodulatory systems during perceptual belief updating and response selection. Leveraging a hierarchical Bayesian learning model to decompose hierarchically-related forms of uncertainty meant that particular processes could be linked to NA, ACh and DA. While manipulating NA and ACh modulated perceptual uncertainty computations, DA receptor antagonism reduced the sensitivity of the motor system to perceptual estimates.

A key benefit of the pharmacological approach used in the present study is that it permitted direct manipulation of the function of three different neuromodulatory systems and comparison of the resulting psychopharmacological effects to a placebo condition. This is relevant given likely functional overlap between the different neuromodulatory systems, as observed here. Indeed, manipulation of a single neuromodulatory system, or use of a single drug, would be agnostic to such an overlap and could make any one effect appear more relevant and specific than it is. The pharmacological approach also meant that it was possible to extend the interpretations of earlier neuroimaging studies (Iglesias et al., 2013; Payzan-LeNestour et al., 2013), from which it is not possible to infer with certainty that activations in particular brain regions, with inhomogeneous cellular compositions, reflect the activity of specific neuromodulatory neurons.

4.5.1 Overlapping, but dissociable, noradrenergic and cholinergic influences on perceptual belief updating

Considerable overlap in the influence of NA and ACh antagonism on perceptual belief updating was identified, but there were also quantitative differences between the drug conditions. While part-synergistic, part-antagonistic interactions between the two neuromodulators during uncertainty processing have been theorised previously (Yu and Dayan, 2005), to my knowledge this is the first study to directly assess these putative computational roles, and to distinguish them from dopaminergic effects, under three pharmacological manipulations and within the same computational framework. I propose that ACh guides probabilistic learning within environmental contexts, while NA has a more circumscribed role in modulating the rate at which an agent learns about the volatility latent in the environment.

4.5.1.1 Noradrenaline influences beliefs about unexpected environmental changes

The present results suggest that NA antagonism under prazosin altered the rate at which individuals updated their volatility beliefs, as indicated by an increase in the model parameter ϑ. An influence of NA- on ϑ fits with the theorised role for NA in computing uncertainty arising from changes in environmental context (Yu and Dayan, 2005). Numerous studies have offered evidence that the NA system is sensitive to highly unexpected events that arise from a hidden contextual change. Noradrenergic neurons in the rat and nonhuman primate LC are responsive to environmental novelty and unexpected changes in reward contingencies (Sara and Segal, 1991; Vankov et al., 1995; Aston-Jones et al., 1997; Bouret and Sara, 2004). Additionally, changes in pupil diameter, attributed at least in part to noradrenergic LC activity (Murphy et al., 2014; Varazzani et al., 2015; Joshi et al., 2016), have been shown to correlate with unexpected outcomes (Preuschoff et al., 2011; Nassar et al., 2012; Browning et al., 2015). I will return to this concept in Chapter 6.

More specifically, in the present study, faster volatility belief updating was observed following NA antagonism. In the HGF model, ϑ represents the volatility of the volatility, and thus the results suggest that NA stabilises an agent's estimate of environmental volatility. This is compatible with the notion that the volatility estimate has a relatively low baseline level, to which it returns after being pushed away. In a volatile environment, this is not an adaptive feature. Rather, the volatility estimate should remain high to enable revision of one's beliefs. It is possible that NA prevents the volatility estimate from falling by reducing an agent's ϑ estimate.

The neurophysiological literature has distinguished two functional modes of LC noradrenergic release (Aston-Jones and Cohen, 2005b; Bouret and Sara, 2005). A phasic mode, characterised by a relatively low baseline firing rate and high phasic responsiveness to task relevant stimuli, has been linked to enhanced task engagement, and a tonic mode to increased distractibility, attention-shifting and exploratory behaviour (Aston-Jones et al., 1994; Usher et al., 1999; Aston-Jones and Cohen, 2005b, but see Jepma et al., 2010). More recently, BOLD activity in the human LC was demonstrated to correlate with "unexpected uncertainty" induced by a switch in reward probabilities associated with familiar stimuli (Payzan-LeNestour et al., 2013), although the negative sign of this correlation still seems to lack explanation. In both my task and that used by Payzan-LeNestour et al., contextual switches required participants to identify discrete changes in underlying transitions between familiar stimuli. To continue making accurate

predictions in light of new transition probabilities, participants had to increase their attentional engagement to facilitate an augmented learning rate. It is likely that in both cases a phasic LC activity mode was recruited, and that this would be recognised as a decrease in BOLD activity at a neuronal population level. Speculatively, it also suggests that the pharmacological NA manipulation in my study may have enabled more phasic NA responsiveness to emerge under suppression of tonic NA firing. Future investigations of the impact of noradrenergic drugs on LC activity profiles are needed to validate this theory.

4.5.1.2 Acetylcholine balances the attribution of uncertainty within and between environmental contexts

Muscarinic ACh receptor antagonism by biperiden led to slower updating of beliefs about stimulus transition contingencies, and so slower adaptation to the probabilistic contexts, as reflected by a decrease in the model parameter ω. I argue that this slowed adaptation also had knock-on effects higher up in the inferential hierarchy. Specifically, I propose that participants attributed perceived violations of their expectations to gross contextual switches as opposed to chance fluctuations in stimulus outcomes, which would be expressed as an increase in ϑ. In light of previous work, which I discuss next, it seems reasonable to suggest that by setting the rate at which an agent learns probabilistic associations, ACh facilitates the appropriate attribution of violated expectations to chance fluctuations in an environment's statistical regularities, or to gross switches in environmental context.

According to the structure of the HGF, a reduction in ω maps onto a reduced precision-weighting of perceptual belief updates at level 2 (compare Equation 3.15 and Equation 3.11). The present findings indicate that under biperiden less weight was given to sensory evidence, and updates of probability estimates became more reliant on current beliefs. This supports proposed roles for ACh in regulating the relative influences of stimulus-driven versus expectation-guided processing (McGaughy et al., 2008; Bentley et al., 2011) and attentional deployment (Bucci et al., 1998; Chiba et al., 1999). For instance, it has been shown that pharmacologically stimulating ACh augments bottom-up sensory signalling in human primary auditory cortex in response to auditory stimuli, possibly by enhancing the gain of superficial pyramidal cells, to bias inference towards sensory data (Moran et al., 2013).

In a recent study, Vossel et al. examined perceptual belief updating during a probabilistic attentional cueing paradigm. By applying a similar instantiation of the HGF to saccadic reaction times, the authors demonstrated faster learning about contextual probabilities

following administration of galantamine, an acetylcholinesterase inhibitor which increases the synaptic availability of ACh, as indicated by an increase in model parameter ω (Vossel et al., 2014a). In the present study, the opposite behavioural effect was observed with the opposite pharmacological manipulation (ACh receptor antagonism), offering independent evidence that ACh signalling guides belief updating about probabilistic associations within environmental contexts.

The present results also indicate that ACh antagonism led individuals to update their volatility estimates more rapidly, reflected by an increase in the model parameter ϑ. This is consistent with the notion that ACh- participants' impaired ability to learn transition contingencies led them to infer that contexts changed at a faster rate. Notably, in their theoretical framework, Yu and Dayan predicted that ACh depletions should cause an agent to underestimate the amount of randomness in a given context. In turn, this causes chance events occurring within a context to seem more significant than they are, meaning they are more likely to be incorrectly taken as indicative of a context change (see Figure 6D in Yu and Dayan, 2005). My experimental observations support this hypothesis and are compatible with data indicating that cholinergic antagonists increase distractibility (Jones and Higgins, 1995) while agonists suppress it (Prendergast et al., 1998; Terry et al., 2002; O'Neill et al., 2003).

It should be noted that, although the perceptual quantities used in this current work are not identical to those previously introduced by Yu and Dayan (Yu and Dayan, 2005), the HGF does embody versions of the same forms of uncertainty. The highest level of uncertainty in the Yu and Dayan (YD) framework was induced by abrupt, discrete, changes in contingencies, which induced what YD call "unexpected uncertainty" (and ascribed to NA). By contrast, the highest level of uncertainty in the HGF is the overall instability of the world, i.e., the rate at which volatility changes. It is this that I found to be modulated by the NA antagonist. Conversely, YD's notion of "expected uncertainty" (ascribed to ACh) suggests that it arises from the known unreliability of predictive relationships within a familiar environmental context. Amongst other effects, the lower the expected uncertainty, the slower the learning – consistent with the effect of parameter ω in the HGF, which was found to decrease under cholinergic antagonism. Along with YD, I also argue that this change in learning has further knock-on effects for what participants perceive to be a chance random event, or a change of context (and hence unexpected uncertainty).

In sum, my findings offer empirical support for the theoretical proposal that ACh and NA interact to construct appropriate cortical representations of volatile contexts, which

facilitates optimal inferences about the current environment (Yu and Dayan, 2005). By regulating high-level uncertainty representations, the two neuromodulators contribute to the updating of an individual's perceptual beliefs, both within and between environmental contexts, an idea that is broadly supported by recent neuroimaging (Iglesias et al., 2013; Payzan-LeNestour et al., 2013; Diaconescu et al., 2017) and pharmacological (Vossel et al., 2014a) evidence.

4.5.2 Dopamine sensitises motor responses to environmental volatility

As per its construction, the present instantiation of the HGF allowed me not only to characterise perceptual belief updating under three pharmacological manipulations, but also to assess how each intervention influenced the deployment of motor responses in in light of individual estimates of uncertainty. Pharmacologically manipulating DA and ACh altered the degree to which participants' perceptual beliefs modulated the preparation of their speeded responses to uncertain stimuli. In contrast, NA antagonism had no significant impact on the sensitivity of participants' motor responses to their current perceptual beliefs, relative to placebo.

I had originally predicted that an individual's capacity to modulate response selection following a sensory PE would be dependent on DA. Indeed, it has previously been shown that pharmacological DA depletion impedes adaptive reactions to unexpected events occurring within predictable contexts (Bestmann et al., 2014). However, in the present study, there was no evidence to suggest that DA receptor antagonism influenced participants' reactions to low-level sensory PE (δ_1). Rather, suppressing DA significantly reduced β_3, which I interpret as a reduction in the sensitivity of participants' motor responses to their higher-level phasic volatility estimates (μ_3).

It is important to note that some key differences distinguish the present experimental design from previous paradigms. In earlier work, participants were pre-trained to respond to stimuli presented within one predictable context, defined by one transition matrix. Furthermore, switches from predictable to unpredictable contexts, consisting of random presentations of stimuli, were explicitly signalled (Bestmann et al., 2014). Therefore, any probabilistic learning and higher-level perceptual uncertainty was removed. In this earlier setting, dopaminergic antagonism under haloperidol selectively impaired participants' reactions to unexpected events that elicited large sensory PEs.

In contrast, the present task created a more complex, and arguably more ecologically valid, scenario in which individuals had to infer the current context for themselves and adapt to any contextual changes. Here, uncertainty representations had to be acquired

through direct sampling from a distribution of observations. To my knowledge, the current study is the first attempt to interrogate the impact of NA, ACh and DA on non-rewarded probabilistic learning within a single behavioural paradigm and a unified Bayesian framework. By estimating beliefs about various forms of uncertainty, I sought to identify neuromodulatory contributions specifically related to particular forms of uncertainty, as opposed to any confounding variables.

Related to this point, a large body of literature examining the role of DA in the context of PE has focused on reward, rather than sensory, PE. Specifically, it is widely thought that phasic activity of dopaminergic neurons in the midbrain signals the discrepancy between the predicted and experienced reward of a particular event (Schultz et al., 1997; Hollerman and Schultz, 1998; O'Doherty et al., 2003; Nakahara et al., 2004; Bayer and Glimcher, 2005; Abler et al., 2006; Daw and Doya, 2006; Pessiglione et al., 2006; D'Ardenne et al., 2008; Hare et al., 2008; Matsumoto and Hikosaka, 2009; Zaghloul et al., 2009; Diederen et al., 2017). The fact that probabilistic learning was unrewarded in the present experiment is one possible reason why no dopaminergic effects on motor responses to low-level sensory PE were observed.

Nonetheless, there have been reports of a role for DA in PE signalling outside the framework of reward (Redgrave et al., 1999; den Ouden et al., 2012; Friston et al., 2012; Galea et al., 2012; Bestmann et al., 2014; Tomassini et al., 2015). Further, it should be noted that the HGF's perceptual model only outputs participant-specific (constant) parameters at the higher levels. As such, in the present study, it was not possible to compare the effects of DA antagonism on sensory PE to Placebo using the approach adopted to examine the effects of NA and ACh on parameters ϑ and ω. Within the framework of the current instantiation of the HGF, it would have been possible to identify an altered effect of sensory PE on motor responses (parameter β_1) under DA antagonism, but not a general effect on the perception of sensory PE. Using fMRI, Iglesias et al. observed that activity in the dopaminergic midbrain correlated with precision-weighted sensory PE (parameter ε_2) during an alternative probabilistic learning task (Iglesias et al., 2013), providing a further indication that DA is involved in updating beliefs in light of low-level sensory PE. Indeed, the authors identified this correlation both when learning was orthogonal to monetary reward and when reward was omitted from the behavioural task entirely. Future work combining neuroimaging and pharmacological manipulations of DA will help to pinpoint the neuromodulator's precise role in perceptual belief updating and response modulation.

The finding that haloperidol reduced the sensitivity of participants' responses to their phasic volatility estimates does sit well with an alternative line of work highlighting the importance of DA in behavioural switching (Cools et al., 2009; van Holstein et al., 2011). For instance, Parkinson's disease patients with DA dysfunction have an impaired capacity to switch from naming digits to letters when both types of stimuli are presented simultaneously, even when the task-shift is explicitly cued (Cools et al., 2001a). In summary, I propose that DA antagonism suppressed response modulation by impeding switching following complex contextual rule changes.

Muscarinic ACh receptor antagonism under biperiden also led to decreased response modulation by parameters at all three hierarchical levels, sensory PE (δ_1), precision-weighted contingency PE (ε_3) and phasic volatility estimates (μ_3), compared to Placebo. I propose that ACh receptor antagonism impeded participants' abilities to learn the statistical structure of the behavioural task, which in turn impaired their capacities to respond accordingly. Although both ACh and DA had effects on response modulation, in light of previous work, I suggest that DA's role is to modulate motor responses according to the widespread perceptual effects of ACh.

4.5.3 Limitations and future work

One of the main constraints of the study is that although prazosin, biperiden and haloperidol are rather selective for NA, ACh and DA receptors respectively, there are complex interactions and dependencies between noradrenergic, cholinergic and dopaminergic systems. Such interactions are a main reason why direct quantitative comparison between drug groups would not have provided direct comparisons between the action of different neuromodulators, and therefore why the current study was designed to detect changes relative to placebo instead. While the results highlight qualitative differences in how NA, ACh and DA influence perceptual belief updating, future work will have to conduct direct quantitative comparisons of their roles.

Further, it is the receptors rather than the neuromodulators themselves that bring about psychophysiological effects, and there are dissociable roles of different receptor sub-types. For instance, the functions of nicotinic versus muscarinic cholinergic receptors in uncertainty signalling have yet to be directly compared. Distinctions have also been made between D1 and D2 dopaminergic receptor sub-types in regulating adaptive responses to unexpected stimuli (Bestmann et al., 2014). Thus, future work could usefully be extended with a range of selective agonists and antagonists for different receptor sub-types. In Chapter 5, I adopt an alternative behavioural genetics approach

to investigate the effects of natural inter-individual variations in DA neurotransmission on perceptual belief updating and response modulation.

Finally, it is likely that all the neuromodulators operate over multiple timescales - for instance, separate, even competing, tonic and phasic effects have been a special target of investigation for NA (Aston-Jones and Cohen, 2005a). Teasing these timescales apart more fully is an ambition for the future, requiring a temporally richer design. Nevertheless, the current findings emphasise the necessity of studying the NA, ACh and DA systems conjointly, as tasks associated with uncertainty will tend to involve them all.

4.5.4 Conclusion

In summary, these results offer novel and direct insight into the complex and intricate effects of NA, ACh and DA during a PSRTT. Employing a hierarchical Bayesian learning model to interrogate various forms of uncertainty and PE, provided interventional evidence linking ACh and NA to uncertainty computations within and between behavioural contexts. In contrast, DA appears to be involved in sensitising motor responses to perceptual volatility estimates. While pharmacological manipulations do not selectively target particular neuromodulatory systems, the results offer a fresh perspective on the effects of noradrenergic, cholinergic and dopaminergic neurotransmission on the computational mechanics of perceptual belief updating according to Bayesian principles. Future studies will verify the generality of the observed effects to different behavioural paradigms with and without learning, reward, prediction and action. By characterising uncertainty computations and response modulation, the methodology reported here could also be used to offer fresh insight into the numerous neurological and psychiatric disorders in which there is dysregulation of processes dependent on NA, ACh and DA.

5 Genetic fingerprints of uncertainty

5.1 Abstract

Behavioural genetics offers an alternative means by which to investigate the relative contribution of different neuromodulators to human learning and action under uncertainty. A range of proteins, from receptors to transporters and degradative enzymes, regulate dopaminergic, noradrenergic and cholinergic neurotransmission. Polymorphisms in the genes that encode these proteins give rise to natural inter-individual variations in neuromodulatory function and to alterations in behaviour. The $Val^{158}Met$ polymorphism in the *COMT* gene has received particular attention within the behaviour genetics literature, with an array of studies having identified variations in dopaminergic neurotransmission and behavioural flexibility as a function of *COMT* genotype. In this chapter, I employ the same probabilistic serial reaction time task (PSRTT) and the same instantiation of the Hierarchical Gaussian Filter (HGF) model used in Chapter 4 to study individual computations of uncertainty and motor response modulation in a naïve sample of 116 healthy human volunteers. I first replicate the behaviour displayed by the Placebo participants in Chapter 4, and verify the capacity of the HGF to capture individual perceptual belief updating and response modulation within a single computational framework of irreducible, estimation and volatility uncertainty. Next, I examine the impact of dopaminergic neurotransmission on these processes by assessing perceptual belief updating and response modulation as a function of *COMT* genotype. The participant sample size is shown to be insufficient to ascertain whether *COMT* genotype has any impact on dopamine-specific learning or action under uncertainty. I discuss how future behavioural genetics approaches could offer fresh insight into the relative roles of dopamine (DA), noradrenaline (NA) and acetylcholine (ACh) to learning and action in dynamic probabilistic environments.

5.2 Introduction

Pharmacological manipulations offer one methodological tool with which to assess the neuromodulatory underpinnings of learning and action in uncertain environments. However, as discussed in Chapter 4, the technique does have its caveats. There are complex interactions and dependencies between different neuromodulatory systems, and different pharmacological agents have different specificities for different receptor sub-types. An alternative approach is to examine learning and action under the natural variations in neuromodulatory function that occur due to polymorphisms in the genes encoding neuromodulatory receptors, transporters and degradative enzymes (Frank et al., 2007, 2009; Tan et al., 2007a, 2007b; Green et al., 2008; Ullsperger, 2010; den Ouden et al., 2013; Doll et al., 2016).

As discusses in Chapter 1, in the context of probing the relative contributions of NA, ACh and DA to learning and action under uncertainty, polymorphisms in the genes that encode the dopamine transporter (DAT), the noradrenaline transporter (NET), the degradative enzymes catechol-O-methyltransferase (COMT) and acetylcholinesterase (ACHE), and the dopaminergic D2-receptor are of particular interest. A summary of the functions of these five proteins, known polymorphisms in the genes that encode them, and any established impact on neuromodulatory phenotype is provided in Table 5.1. For additional details, please refer back to Chapter 1.

Gene	Function	Polymorphism	Phenotype
COMT	Encodes the COMT enzyme, which catalyses the degradation of catecholamines, including DA, especially in the prefrontal cortex	$Val^{158}Met$ single nucleotide polymorphism (SNP) at rs4680, resulting in Val and Met alleles	Met allele is associated with decreased COMT activity and increased DA neurotransmission
DAT1	Encodes the DAT, which mediates the reuptake of DA from the synaptic cleft, especially in the striatum	Variable number tandem repeat (VNTR) at rs28363170 (the 3' untranslated region), commonly resulting in 9- (9R) and 10-repeat (10R) alleles	9R allele has been associated with altered DAT availability and a putative change in DA neurotransmission, but the precise functional impact is speculative

DRD2	Encodes the DA D2-receptor, of which there is a particularly high striatal density	SNP at rs1800497, resulting in A1 and A2 alleles	A2 allele associated with increased DA D2-receptor expression
NET	Encodes the NET, which mediates the reuptake of NA from the synaptic cleft	SNP at rs2242446, resulting in C and T alleles	Any functional impact on NA neurotransmission is unclear
ACHE	Encodes the ACHE enzyme, which catalyses the degradation of ACh	SNP at rs2571598, resulting in A and G alleles	Any functional impact on ACh neurotransmission is unclear

Table 5.1 Summary of genetic polymorphisms that modulate neuromodulatory function. *The Val^{158}Met SNP in the COMT gene is one of the best studied polymorphisms in the behavioural genetics literature, with effects on COMT activity and DA neurotransmission being relatively well established. Polymorphisms in the DAT1 and DRD2 genes are also thought to impact on DA neurotransmission, but their precise functional effects are speculative. Any functional impact of the NET and ACHE polymorphisms on neuromodulatory transmission is also currently unclear.*

Behavioural genetics has several methodological advantages. First, it permits the effects of different neuromodulatory systems to be assessed within individuals and in a single experimental session. Second, it offers a means by which to investigate the relative contribution of neuromodulators to learning and action without any confounding effects of pharmacological interventions. For instance, pharmacological agents are often not wholly specific for particular receptor sub-types and they likely modify the baseline dynamics, interactions and compensatory mechanisms of different functionally-coupled neuromodulatory systems. Third, identifying the functional consequences of polymorphisms in the genes that encode different neuromodulatory receptors and transporters, with different relative distributions throughout the brain, holds the potential to better elucidate the contributions of different neuromodulatory signalling pathways to learning and action under uncertainty.

This chapter was motivated by the possibility to characterise the effects of *COMT*, *DAT1*, *DRD2*, *NET* and *ACHE* genotypes on learning and response modulation within

individuals undertaking the same PSRTT applied in Chapter 4. To recap, the PSRTT exposes participants to three distinct forms of uncertainty: *irreducible uncertainty* arising from the inherent randomness of the probabilistic transitions between consecutive stimuli, *estimation uncertainty* arising from an individual's imperfect knowledge of the probabilistic relationships governing stimulus transition contingencies within contexts, and *volatility uncertainty* arising from contextual instability. By applying the novel instantiation of the HGF model to the behavioural data, the original aim was to characterise the impact of dopaminergic, noradrenergic and cholinergic genotypes on perceptual belief updating and response modulation in dynamic probabilistic environments.

While the influence of *COMT* genotype on cortical dopaminergic neurotransmission is relatively well established (Gogos et al., 1998; Männistö and Kaakkola, 1999; Akil et al., 2003; Chen et al., 2004; Tunbridge et al., 2004; Yavich et al., 2007), investigations of learning and action as a function of *DAT1*, *DRD2*, and particularly *NET* and *ACHE,* genotypes would have been more exploratory. Indeed, the functional impact of the latter four genes on dopaminergic, noradrenergic and cholinergic neurotransmission is currently more elusive. Nevertheless, behavioural investigations of polymorphisms in the *COMT*, *DAT1*, *DRD2*, *NET* and *ACHE* genes hold the potential to extend the pharmacological results in Chapter 4 by further elucidating the roles of DA, NA and ACh in learning and response modulation under uncertainty. In particular, in light of the finding that pharmacological DA antagonism reduced the sensitivity of motor responses to phasic volatility estimates but not to sensory prediction error (PE), a key motivation for investigating any impact of *COMT* and *DAT1* genotypes on response modulation within the same computational framework was the possibility to identify separable cognitive (cortical) and motoric (striatal) DA-mediated processes underlying flexible behaviour, respectively.

Unfortunately, careful scrutiny of the genetic data revealed that the laboratory that undertook the genotyping analyses had not provided reliable genotypic summaries for all five genes in the first cohort of individuals from whom genetic data was collected. As such, participant recruitment was halted early. One gene for which I do have reliable genotypic data from 116 participants is *COMT*. Therefore, in the following, I focus on assessing the effects of three dopaminergic *COMT* genotypes on perceptual belief updating and response modulation during the PSRTT. This approach complemented the methodology employed in Chapter 4 by facilitating an alternative examination of dopaminergic contributions to learning and action in uncertain environments, focusing

on cortical DA neurotransmission and free from potentially confounding effects of pharmacological DA manipulations.

5.2.1 The Val^{158}Met *COMT* polymorphism

To reiterate Chapter 1, the *COMT* gene encodes the COMT enzyme which catalyses the degradation of catecholamines, particularly cortical DA (Gogos et al., 1998; Akil et al., 2003; Tunbridge et al., 2004; Yavich et al., 2007). A single nucleotide polymorphism (SNP) at rs4680 results in an amino acid switch from valine (Val) to methionine (Met), at codon 158. The Met isoform has reduced thermostability at body temperature, resulting in a 3-4 fold decrease in COMT enzymatic activity compared to the Val isoform, and so higher synaptic DA concentrations (Männistö and Kaakkola, 1999; Chen et al., 2004). In contrast, the Val allele is associated with higher enzymatic activity and so lower synaptic DA availability.

5.2.2 Probing a role for dopamine in perceptual belief updating and response modulation

In Chapter 4, dopaminergic antagonism under haloperidol was found to decrease the sensitivity of participants' motor responses to their beliefs about the environment's volatility. In contrast, no effects of DA antagonism were observed on the rate at which participants learned about contextual transition contingencies or the volatility of these contingencies over time. As discussed in Chapter 4, I had originally predicted that DA would modulate an individual's capacity to modulate response selection following a low-level sensory PE owing to previous work demonstrating that pharmacological DA depletion impedes adaptive reactions to unexpected events occurring within predictable contexts (Bestmann et al., 2014). While the absence of an effect of DA on response modulation by sensory PE can be explained by differences in experimental paradigms, and while the finding that DA sensitises motor responses to phasic volatility estimates sits well with a hypothesised role for DA in behavioural switching (Cools et al., 2001a, 2009; van Holstein et al., 2011), in the present experiment I sought to extend these findings beyond a pharmacological approach.

5.2.2.1 Replication of learning and action during the PSRTT

In particular, employing the same PSRTT and novel instantiation of the HGF applied in Chapter 4, I tracked human learning and response modulation in a dynamic probabilistic environment that gave rise to irreducible, estimation and volatility uncertainty. It was therefore possible to investigate whether the learning and behaviour observed in Chapter

4's Placebo group could be replicated in a naïve cohort of healthy individuals, and whether the HGF model would perform as well as it had done in its first application.

5.2.2.2 Replication of dopaminergic effects on learning and action

Next, since the HGF captures an individual's learning of the task's structure and maps their beliefs onto their observed reaction time (RT) responses, participant-specific perceptual belief updating could be disentangled from the sensitivity of motor response to perceptual estimates. Further, it was possible to assess belief updating and response modulation as a function of *COMT* genotype. Given that COMT regulates DA neurotransmission, I aimed to:

1. Replicate the finding that DA sensitises an individual's motor responses to their phasic volatility estimates. Given that Met carriers show lower COMT activity and thus higher DA neurotransmission than Val/Val homozygotes, I hypothesised that motor responses in Val/Met and Met/Met individuals would show increased sensitivity to phasic volatility estimates (Figure 5.1E).

2. Probe whether the sensitivity of an individual's motor responses to their sensory PE varies as a function of *COMT* genotype, suggesting a modulatory role for DA. Based on previous work (Bestmann et al., 2014), I hypothesised that motor responses in Met carriers might show increased sensitivity to sensory PE (Figure 5.1D). However, the results of Chapter 4 would predict no effect of DA on this parameter.

3. Replicate the finding from Chapter 4 that reduced DA neurotransmission leads to general RT slowing, echoing bradykinesia in Parkinson's disease, a disorder characterised by DA depletion in the substantia nigra (Berardelli et al., 2001). I hypothesised that the constant component of RT would be increased in Val/Val homozygotes (Figure 5.1C).

4. Replicate the finding that DA neurotransmission does not modulate the speed at which individuals update their beliefs about phasic volatility (Figure 5.1A) or contextual transition contingencies (Figure 5.1B).

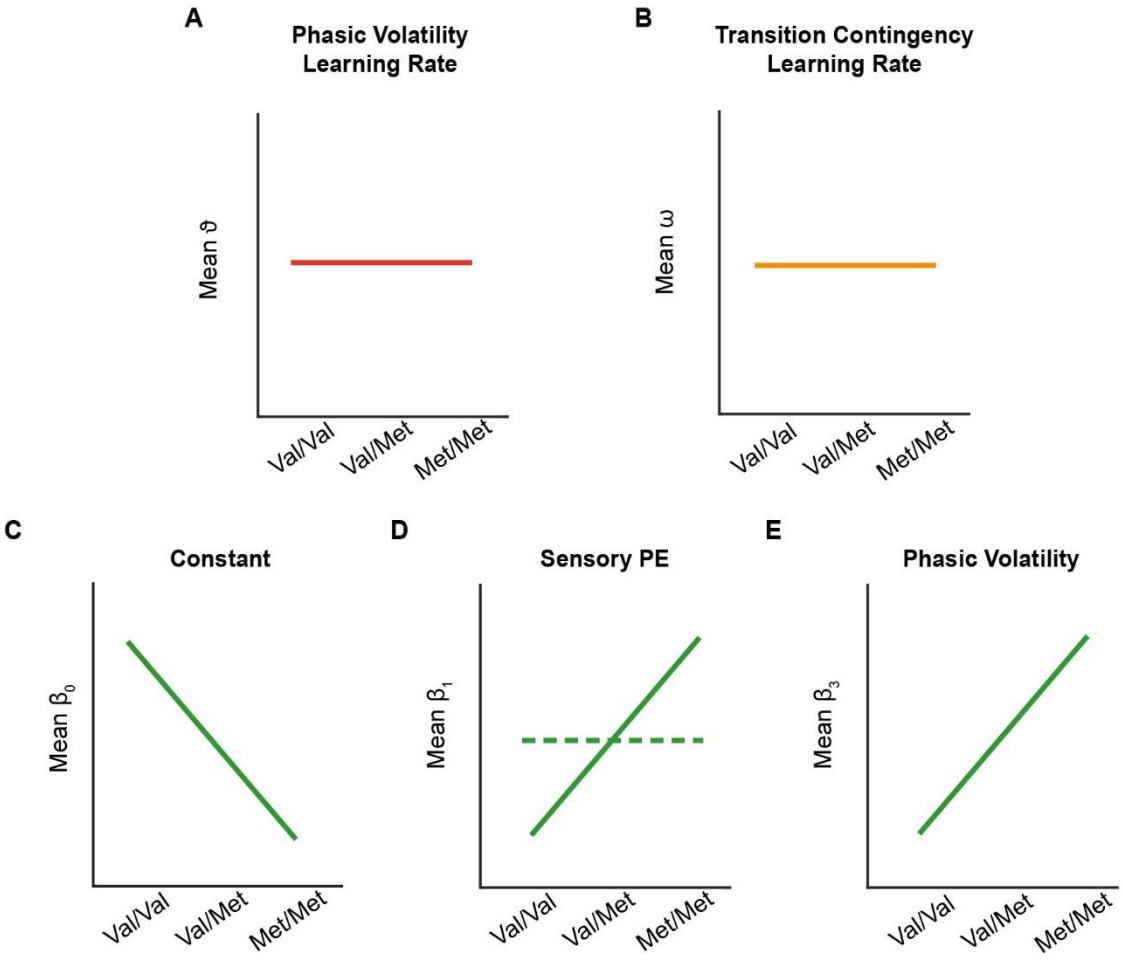

Figure 5.1 Predicted effects of COMT genotype on perceptual belief updating and response modulation. *(A-B) Based on the pharmacological DA results of Chapter 4, increasing DA neurotransmission in Met carriers would be expected to have no effects on perceptual belief updating (i.e., on parameters ϑ and ω). (C-E) The pharmacological findings predict motor response modulation would vary between Met carriers and Val/Val homozygotes. Specifically, increased DA neurotransmission in Met carriers would be expected to decrease the constant component of log(RT) responses (β_0) and increase the sensitivity of motor responses to phasic volatility estimates (β_3). The pharmacological results would predict no effect of COMT genotype on the sensitivity of motor responses to sensory PE (β_1) (D; dashed line) but, given previous work demonstrating that DA depletion is associated with impaired reactions to unexpected events occurring in predictable contexts, I hypothesised that motor responses in Met carriers would show increased sensitivity to sensory PE (D; solid line).*

5.3 Methods

5.3.1 Participants

116 healthy participants (18 male, aged 18-31 years, 73 Caucasian) with normal or corrected-to-normal vision took part in this study after giving written informed consent. The experiment was run in collaboration with the University of Birmingham. The experimental protocol was approved by the University of Birmingham Research Ethics Committee.

5.3.2 Probabilistic serial reaction time task

The experimental setup and PSRTT used were based on those described in Chapter 4. In brief, participants sat facing a computer screen positioned approximately 60cm away. They were instructed to rest their left and right index and middle fingers on four marked keys on a computer keyboard, and to maintain this position throughout the task. On each trial, participants were required to respond to the presentation of one of four visual stimuli by making a speeded button-press before the end of a 1200ms intertrial interval (ITI). Each stimulus was associated with one particular button. The stimulus-response mappings remained consistent within an experimental session but were counterbalanced across participants.

5.3.2.1 Training

Each participant acquired the stimulus-response mappings for their session during a training block in which they received visual error feedback after each trial. The training session comprised at least 100 trials and did not finish until the participant had reached a minimum performance criterion of 85% accuracy on the last 20 trials. Participants were then given 15 practice trials, in which the stimuli were presented in a random order and without error feedback, to familiarise them with the timings of the main experiment. On average, participants responded correctly on 90.4 ± 1.1% (± SEM) of the practice trials, indicating adequate learning and retention of the mappings.

5.3.2.2 Task design

Each participant performed 800 trials of the PSRTT. At any given time, there was an underlying probabilistic rule, defined by one of eight transition matrices (TMs), which determined the probabilistic relationship between the stimulus presented on trial, *t*, and the stimulus presented on the previous trial, *t-1*. The TM switched every 50 trials without explicit indication to the participant. The TMs comprised the two 1^{st}-order and four 0^{th}-order TMs used in Chapter 4 (Figure 2.1), as well as two additional 1^{st}-order TMs (Figure

2.2). Trials were drawn from each TM twice. The order of TMs was pseudorandom, with no consecutive repeats. The overall probability of each stimulus was equal across the 800 trials.

As in Chapter 4, the different TMs created contexts that the participants could infer from stimulus observations. For fast and accurate responses, participants had to track *irreducible uncertainty* arising from the inherent randomness of the probabilistic transitions between consecutive stimuli; *estimation uncertainty* arising from their imperfect knowledge of the probabilistic relationships governing stimulus transition contingencies within contexts; and *volatility uncertainty* arising from the unsignalled contextual instability.

The pseudorandom order of TMs was used to generate one stimulus sequence that was used for all participants to ensure comparable learning processes and model parameter estimates. Rest periods occurred every 215 trials, orthogonal to TM switches. The importance of fast responses was stressed. Participants were told that by paying attention to any patterns in the order in which stimuli were presented, and to any switches in these patterns, it may be possible to respond faster. No further information about the nature of the experiment was provided. Anticipatory responses (<80ms) were recorded as incorrect.

5.3.3 General procedure

Participants were recruited from the University of Birmingham Undergraduate Psychology Student cohort. To ascertain the effects of DA neurotransmission on learning and action in uncertain environments, any differences in perceptual belief updating and response modulation during the PSRTT were assessed as a function of *COMT* genotype.

5.3.4 Genotyping

Genomic DNA was extracted from saliva samples collected from each participant using the Oragene OG-500 self-collection kit (Oragene, DNA Genotek Inc., Canada) according to the manufacturer's recommendations. Molecular genetic analyses were performed at the West Midlands Genetic Laboratory at Birmingham Women's Hospital. Next generation sequencing was conducted to genotype the SNP at rs4680 within the *COMT* gene.

Salivary DNA was extracted according to standardised protocols and quantified using Qubit fluoriometry. Genotyping was carried out by multiplex polymerase chain reaction (PCR) amplicon resequencing. In brief, PCR amplicons flanking the SNP of interest were

designed via identification of their position in the UCSC hg19 reference genome. Reference sequences for +/- 200 base pairs flanking the SNP were retrieved and a sequencing amplicon was designed using Primer 3 (Untergasser et al., 2012) set for an annealing temperature of 60^{o}C and with design conditions as recommended for Fluidigm Access Array primer design. Primers were validated using gradient PCR. 20ng of genomic DNA collected from participants was pooled and barcoded. Multiplex amplicon PCR was carried out using the Fluidigm Access Array system using standard conditions. The multiplexed library was diluted to a loading concentration of 4pM and sequenced on an Illumina MiSeq using a v2 500 cycle kit. Each amplicon was sequenced to a minimum read depth of 2000x. FASTQ files were exported from the sequencing instrument, quality trimmed with Trimgalore and aligned to the hg19 reference genome using Bowtie2 (Langmead and Salzberg, 2012). SNP variants were called using FreeBayes (Garrison and Marth, 2012) and annotated using ANNOVAR (Yang and Wang, 2015). Variant QC was carried out by visual inspection of 10% of amplicon calls for 10% of samples using the UCSC genome browser.

PCR primers were designed to flank the SNP, producing a 249 base pair amplification product (Table 5.2).

	Forward primer	**Reverse primer**
COMT	CGAGGCTCATCACCATCGAG	GGGAGGACAAAGTGCGCAT

Table 5.2 Sequence primers for the COMT SNP (rs4680).

5.3.5 Model-agnostic analyses

Trial-wise RT was calculated as the time between stimulus onset and the subsequent button press. The RT data were log-transformed (Bestmann et al., 2014; Marshall et al., 2016). A series of conventional, model-agnostic analyses of behaviour were first conducted to assess whether participants learned about the underlying stimulus transition contingencies, whether the behavioural data replicated that observed in the Placebo group in Chapter 4, and whether learning was modulated by *COMT* genotype. To assess the interaction between stimulus transition probability and drug, trials were binned according to three probability levels corresponding to the presented stimuli's true transition probabilities as existed in the TMs (High: 0.85 and 0.70; Mid: 0.20; Low: 0.05) (Galea et al., 2012; Bestmann et al., 2014; Marshall et al., 2016). A repeated-measures analysis of variance (RM-ANOVA) was used to compare mean log(RTs) for correct responses across these three probability levels.

To obtain a model-agnostic indication of learning across the course of the probabilistic contexts, a median split was performed on each 50-trial contextual block. A RM-ANOVA was used to compare mean Δlog(RTs) on correct Early (1-25) vs Late (26-50) trials at each probability level.

To identify any evidence of post-error slowing during the PSRTT, i.e. slower responses on trials following those on which participants made an error (Rabbitt, 1966; Botvinick et al., 2001; Gehring and Fencsik, 2001; Cavanagh et al., 2014; Marshall et al., 2016), a RM-ANOVA was used to compare log(RTs) on correct trials that immediately followed correct and erroneous responses. A further RM-ANOVA compared log(RTs) on correct, post-infrequent trials, i.e., trials following those with a true transition probability of 0.05, and correct trials following trials with a true transition probability >0.05.

5.3.6 Model-based analyses

The same instantiation of the HGF used in Chapter 4 (see Figure 3.4, Figure 4.2 and Figure 5.2) was applied to the behavioural data. To recap, this version comprises a three-level perceptual model and a response model.

5.3.6.1 Perceptual model

The perceptual model tracks each participant's estimated beliefs about the PSRTT's trial-wise stimulus transitions, transition contingencies, the volatility of the transition contingencies, and the respective *irreducible*, *estimation* and *volatility uncertainty* about these beliefs. The participant-specific parameters ϑ and ω capture the respective rates at which an individual updates their beliefs about phasic volatility and transition contingencies, and allows for individual expression of approximate Bayes-optimal learning.

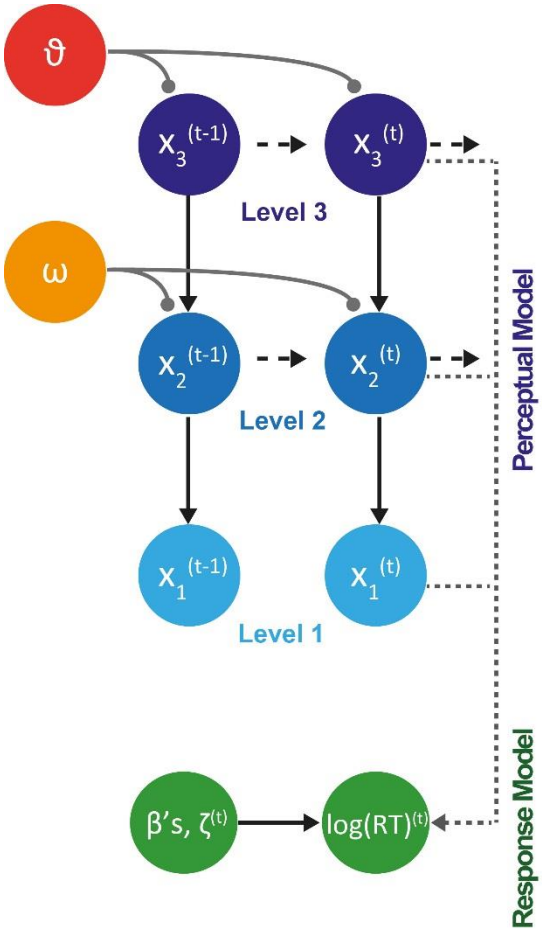

Figure 5.2 The Hierarchical Gaussian Filter (HGF). *The perceptual model tracks an individual's learning across three levels. State* $\boldsymbol{x}_1$ *represents trial-wise stimulus transitions from one stimulus to the next,* $\boldsymbol{x}_2$ *the transition contingencies, and* x_3 *the phasic volatility, where t is the current trial number and bold font is used to indicate a matrix. Participants hold and update beliefs about the true quantities at each level, with a mean μ and a variance σ. ϑ and ω are participant-specific parameters that couple the levels and determine the respective speed of belief updating about phasic volatility and transition contingencies. The response model describes the mapping from a participant's trial-wise beliefs onto their observed log(RT) responses.*

5.3.6.2 Response model

The response model provides a mapping from each participant's trial-wise beliefs, as provided by the perceptual model, onto his/her observed log(RT) responses. To verify that the response model applied in Chapter 4 was equally applicable in the present study, three response models were constructed and compared using random effects Bayesian model selection (Stephan et al., 2009; Rigoux et al., 2014). The three response models were identical to those compared in Chapter 4 (see section 4.3.5.2). To recap, the first specified that trial-wise log(RT) was a linear function of a constant component of log(RT),

sensory PE (δ_1), precision-weighted contingency PE (ε_3), estimated phasic volatility (μ_3), post-error slowing, and Gaussian noise (ζ).

Response Model 1:

$$\log(\mathrm{RT})^{(t)} = \beta_0 + \beta_1(\delta_1^{(t)}) + \beta_2(\varepsilon_3^{(t)}) + \beta_3(\mu_3^{(t)}) + \beta_4(\mathrm{PostError}^{(t)}) + \zeta^{(t)}$$

Equation 5.1

The second contained the precision-weighted form of sensory PE (ε_2) instead of δ_1.

Response Model 2:

$$\log(\mathrm{RT})^{(t)} = \beta_0 + \beta_1(\varepsilon_2^{(t)}) + \beta_2(\varepsilon_3^{(t)}) + \beta_3(\mu_3^{(t)}) + \beta_4(\mathrm{PostError}^{(t)}) + \zeta^{(t)}$$

Equation 5.2

Since δ_1 and ε_2 are highly correlated, a third response model containing both parameters was constructed.

Response Model 3:

$$\log(\mathrm{RT})^{(t)} = \beta_0 + \beta_1(\delta_1^{(t)}) + \beta_2(\varepsilon_2^{(t)}) + \beta_3(\varepsilon_3^{(t)}) + \beta_4(\mu_3^{(t)}) + \beta_5(\mathrm{PostError}^{(t)}) + \zeta^{(t)}$$

Equation 5.3

5.3.6.3 Model fitting

The perceptual and response model priors were identical to those used in Chapter 4 (see Table 4.1). As in Chapter 4, the perceptual model assumed that participants updated their beliefs according to the stimulus presented on each trial, while the response model incorporated correct trials only.

5.3.6.4 Parameters of interest

To probe any dopaminergic effects on perceptual belief updating, the participant-specific phasic volatility learning rate (ϑ) and transition contingency learning rate (ω) were assessed as a function of *COMT* genotype. Similarly, to identify any dopaminergic effects on response modulation, the sensitivity of participants' log(RTs) to their sensory PE (β_1) and phasic volatility estimates (β_3) were compared across *COMT* genotypes, as was the constant component of log(RT) (β_0).

5.3.7 Statistical analyses

In reporting statistical differences, a significance threshold of α=0.05 was used. Where assumptions of sphericity were violated (Mauchly's test $p<0.05$), the Greenhouse-Geisser correction was applied.

For comparisons across repeated-measures and across the three *COMT* genotypes (Val/Val, Val/Met and Met/Met), partial eta-squared (η_p^2) is reported as the effect size. For analyses of the perceptual and response model parameters of interest, one-way ANOVAs were first conducted to assess any impact of the three *COMT* genotypes. Due to the different sample sizes across genotypes (Table 5.4), further exploratory analyses were conducted. Here individuals with a Val/Met or Met/Met genotype were grouped to form a single set of 37 Met carriers. Independent t-tests were then applied to compare Met carriers to Val/Val homozygotes. To recap, the Met isoform produces a less active form of COMT, resulting in higher dopaminergic neurotransmission due to reduced catecholamine degradation. Before conducting each independent t-test, Levene's test was used to verify that there was no significant difference in the variances of the populations from which the data samples had been drawn. For independent t-tests, Cohen's *d* is reported as the effect size.

5.3.8 Control analyses

5.3.8.1 Model parameter correlations

To demonstrate that the HGF provided a good fit to the behavioural data, the correlations between the Bayesian parameter averages (BPAs) for the model parameters were assessed.

5.4 Results

Data from 116 participants are reported. A summary of demographics is provided in Table 5.3.

	Participants (n = 116)
Gender	
(number male)	18
(number female)	98
Age ***(years)***	19.6 ± 2.1
Education Level ***(1-5)***	2.2 ± 0.1
Ethnicity (%)	
Caucasian	62.9
Asian	23.3
African	6.0
Mixed Race	6.0
Other	1.7

Table 5.3 Summary details for all 116 participants. *Education Level refers to the highest attained from the following: 1 = compulsory education (≤ 12 years); 2 = further education (13-14 years); 3 = undergraduate degree (15-17 years); 4 = one postgraduate degree (≥ 18 years); 5 = multiple postgraduate degrees. Age data are mean ± SD. Education data are mean ± SEM.*

5.4.1 Model-agnostic results

On average, participants made correct responses on 88.5 ± 0.5% (± SEM) of trials, which was equivalent to the Placebo group's mean correct response rate (90.3 ± 0.8% of trials) in Chapter 4 (t_{146}=1.65, p=0.101). Note that Levene's test indicated that that there was no significant difference in the variances of the populations from which the genetic and Placebo data samples were drawn (F=2.88, p=0.092).

A 3 probability RM-ANOVA conducted on the log(RTs) for correct trials binned according to the four true conditional probabilities that existed in each of the transition matrices, grouped into High (0.85 and 0.70), Mid (0.20) and Low (0.05) transition probabilities, indicated that there was a significant increase in log(RTs) with decreasing transition probability (main effect of probability: $F_{1.63,187.06}$=567.72, p<0.001, effect size η_p^2=0.83; Figure 5.3A).

Moreover, a RM-ANOVA on Δlog(RTs) for Late vs Early trials indicated that, across the course of a contextual block (Figure 5.3B), participants became faster at responding to High and Mid probability and slower at responding to Low probability stimuli ($F_{1.89,217.30}$=132.45, p<0.001, η_p^2=0.54). Together, these results demonstrate that participants showed learning of the true stimulus transition contingencies.

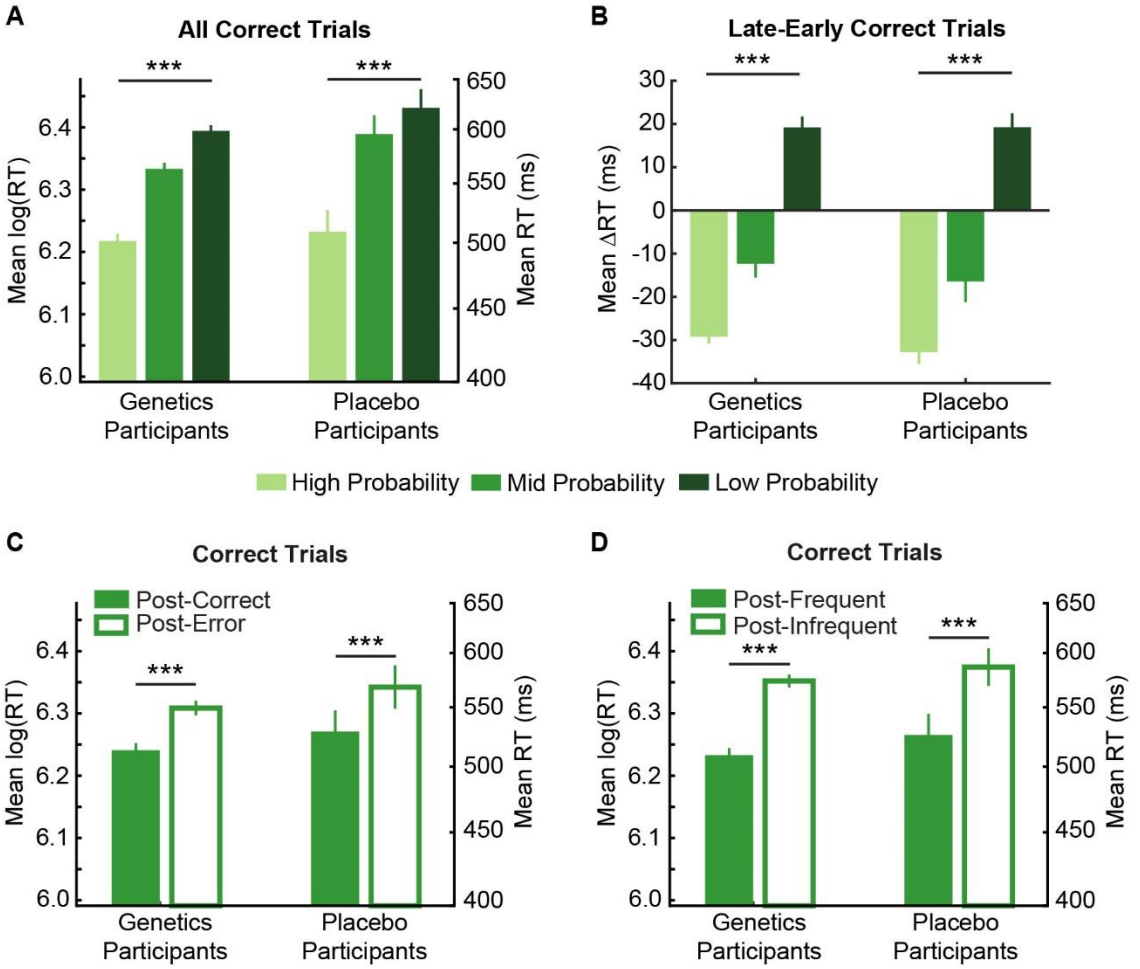

Figure 5.3 Model-agnostic results. *Changes in log(RT) indicate that participants learned to predict the stimulus transitions, echoing the results for the Placebo group in Chapter 4. (A) An increase in log(RT) occurred as a stimulus' true transition probability decreased. (B) A median split on each 50-trial contextual block was used to compare mean log(RTs) on Early (1-25) and Late (26-50) trials at each probability level. Over the course of a context, participants became faster at responding to High and Mid probability stimuli, and slower at responding to Low probability stimuli. Raw RTs are plotted here to simplify interpretation of ΔRT, but statistics were conducted on log(RTs). (C) Across drug-groups, participants showed evidence of post-error slowing on correct trials that followed an erroneous response compared to those that followed correct responses. (D) Participants also showed evidence of slowing on correct trials that followed an infrequent stimulus transition. Results are mean ± SEM. *** p<0.001.*

Participants in the genetics cohort also showed evidence of post-error slowing on correct trials that followed those on which they made an error ($F_{1,115}$=119.14, p<0.001, η_p^2=0.51; Figure 5.3C). Participants also demonstrated significant log(RT) slowing on correct, post-

infrequent trials (true transition probability = 0.05) compared to all other correct trials ($F_{1,115}$=952.43, p<0.001, η_p^2=0.89; Figure 5.3D).

These four findings replicate the results for the Placebo group in Chapter 4. Indeed, repeating the analyses with group (i.e., Genetics cohort or Placebo cohort) as a between-subjects factor revealed no significant between-subjects effect of group on log(RTs) (all p≥0.14).

Repeating each of the analyses with *COMT* genotype (Val/Val, Val/Met, Met/Met) as a between-subjects factor revealed no significant effects of genotype on log(RTs) across true transition probability levels (p=0.680), across the course of contextual blocks (p=0.167), on post-error trials (p=0.082) or on post-infrequent trials (p=0.494).

5.4.2 Model-based results

5.4.2.1 Perceptual model

Overall, the HGF tracked the true stimulus transitions well (Figure 5.4). Note again that the model is uninformed of the true stimulus transition probabilities, but rather bases its estimates on the observed stimulus transitions only.

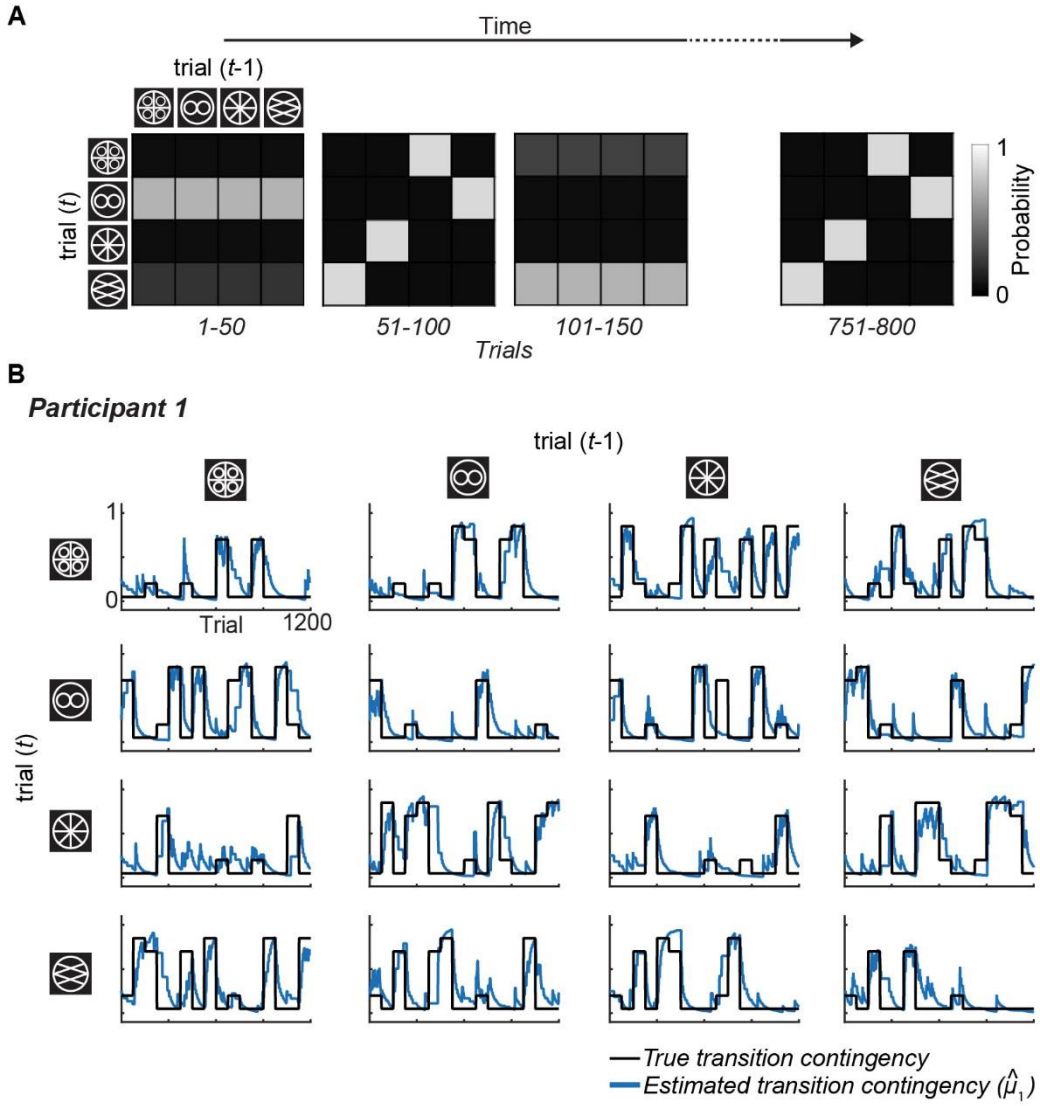

Figure 5.4 Estimated transition contingencies for an example participant. *(A) Transitions between pairs of stimuli, from trial t-1 to trial t, were defined by transition matrices. Every 50 trials the transition matrix switched to a different matrix. (B) Each panel corresponds to one of the 16 possible transitions between stimuli across 800 trials. The black lines indicate the true transition contingencies. The blue lines reflect the participant's inferred estimates (i.e., their posterior expectation of these contingencies, $\hat{\mu}_1$) before seeing the stimulus outcome on each trial. The model tracked the true underlying contingencies and detected change-points. Here, in a representational participant from the Genetics cohort, the model tracked the true transition contingencies closely.*

Again, as demonstrated in Chapter 4, when trials were categorised according to participants' trial-wise estimates of transition contingencies, as provided by model parameter $\hat{\mu}_1$ (five bins: 0.8-1, 0.6-0.8, 0.4-0.6, 0.2-0.4, 0-0.2), the same increase in

log(RT) with decreasing transition probability found in the model-agnostic results was observed (c.f. Figure 5.5 with Figure 5.3A); significant effect of $\hat{\mu}_1$: $F_{2.04,175.53}$=269.14, p<0.001, η_p^2=0.76).

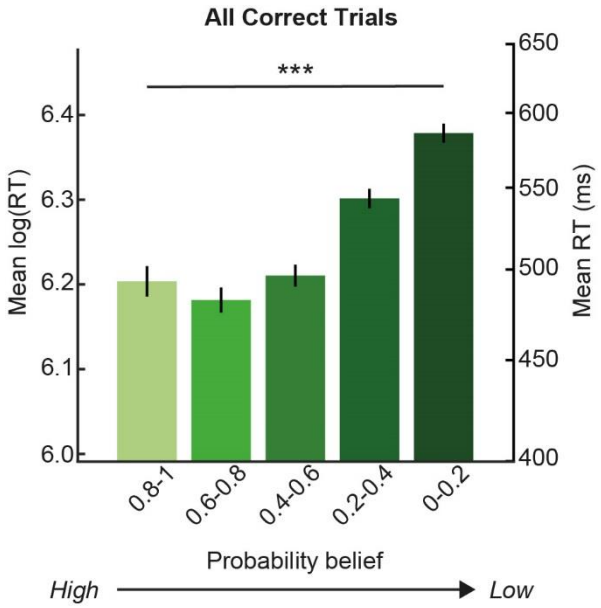

Figure 5.5 Model-based changes in log(RT) mirror the model-agnostic results. *Faster responses were observed as participants' estimates of the true transition contingencies increased, demonstrating that the HGF captured the same behavioural effect identified in the model-agnostic analyses, i.e., that participants learned to predict the stimulus transitions and prepared motor responses to high probability transitions (c.f. Figure 5.3A). Results are mean ± SEM. *** p<0.001.*

5.4.2.2 Response model

Random effects Bayesian model selection established that Response Model 1 (containing parameters δ_1, ε_3 and μ_3) was superior by a considerable margin (posterior probability: 0.5643; protected exceedance probability, i.e., the probability that Response Model 1 is more likely than any other model in the comparison set: 0.9493; Figure 5.6A). This replicates the finding in Chapter 4 that Response Model 1 was superior. Moreover, no significant difference was found in the noise parameter ζ between the Genetics and Placebo cohorts (t_{146}=1.46, p=0.146; Figure 5.6B), indicating that Response Model 1's ability to predict log(RTs) was unaltered across the two experiments. Note that Levene's test established that there was no difference in the variances of the populations from which the Genetics and Placebo ζ samples were drawn (F=1.783, p=0.184).

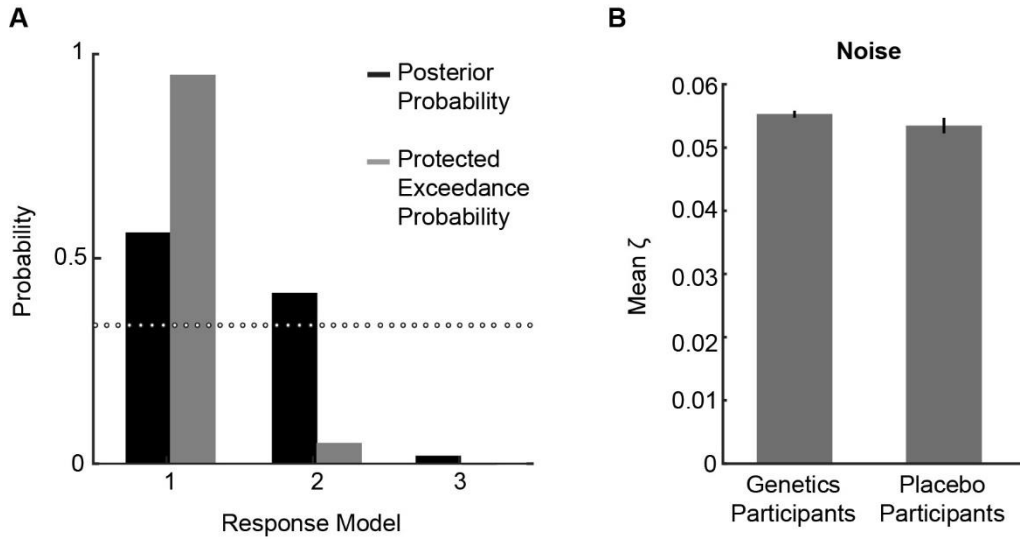

Figure 5.6 Model comparison results. *(A) Random effects Bayesian model selection indicated that Response Model 1 was superior. Posterior probabilities quantify the likelihood of each model given the data. Protected exceedance probabilities quantify how likely it is that any given model is more frequent than all other models in the comparison set while also protecting against the possibility that the observed variability in (log-) model evidence could be due to chance. The dotted line indicates the threshold for chance-level posterior probabilities (p=0.33). (B) The lack of a difference in the noise parameter ζ between the Genetics cohort and the Placebo cohort (p=0.146) indicates that the model's ability to predict log(RT) was unaltered across groups.*

5.4.3 Assessment of the effects of *COMT* genotype on perceptual belief updating and response modulation

The number of participants with each *COMT* genotype is summarised in Table 5.4.

Polymorphism	Genotype	Participants
COMT	Val/Val	79
(rs4680)	Val/Met	24
	Met/Met	13

Table 5.4 Number of participants with each COMT genotype.

5.4.3.2 No identifiable effects of COMT genotype on perceptual belief updating

The rate at which individuals updated their volatility estimates, as reflected by parameter ϑ, was equivalent across the three *COMT* genotypes ($F_{2,113}$=0.76, p=0.471; Figure 5.7A). Similarly, the rate at which individuals updated their transition contingency estimates and thus adapted to the probabilistic contexts, as reflected by parameter ω, was unaltered

by *COMT* genotype ($F_{2,113}$=1.40 p=0.251; Figure 5.7B). These results echo the finding in Chapter 4 that D1/D2 receptor antagonism under haloperidol did not influence the rate at which participants learned about the task's volatility or contextual transition contingencies compared to Placebo.

Similarly, comparing these two perceptual model parameters in Val/Val homozygotes and in individuals with higher DA neurotransmission (Met carriers) also indicated no effect of *COMT* genotype on ϑ (t_{114}=0.50, p=0.618) or ω (t_{114}=-0.84, p=0.404). Note that Levene's test confirmed equality of variances across groups for both ϑ (F=0.72, p=0.397) and ω (F=1.80, p=0.182).

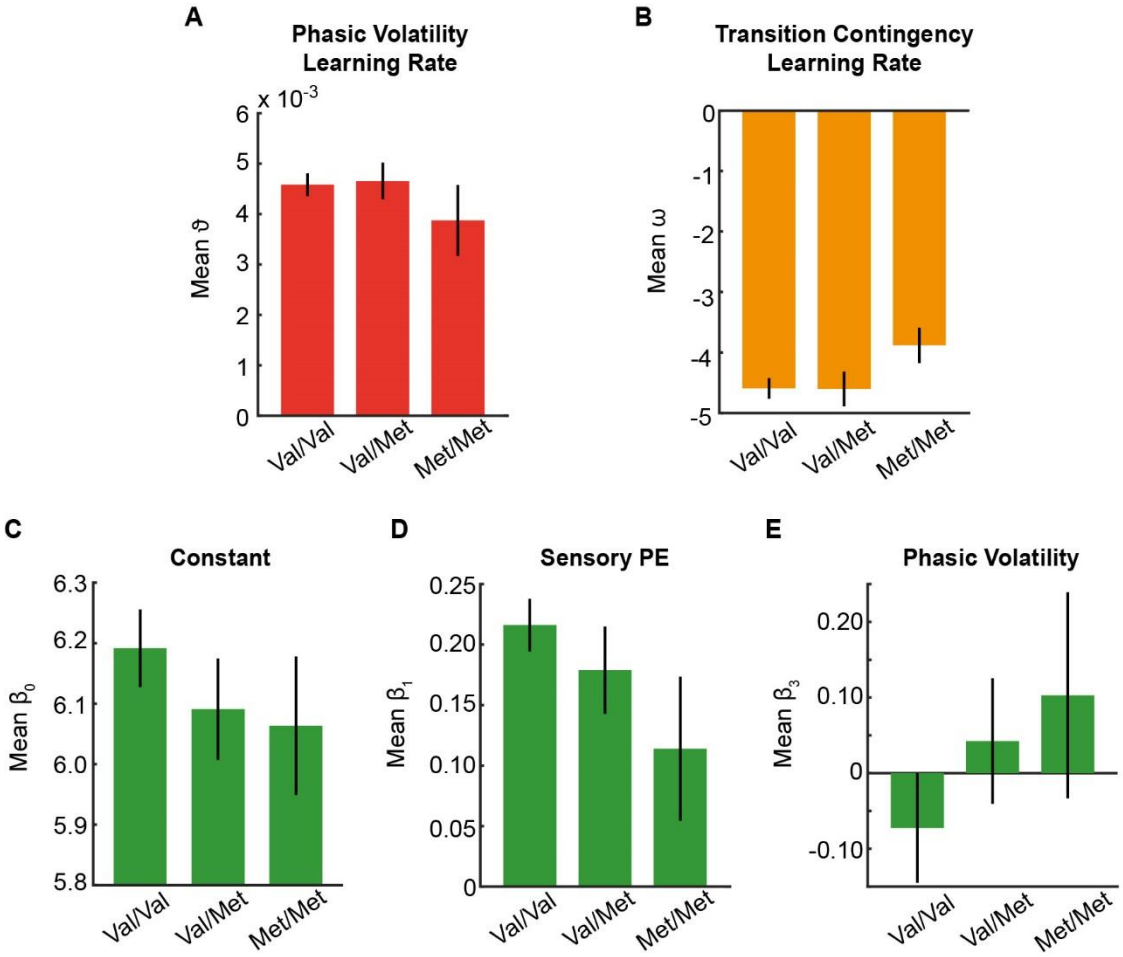

Figure 5.7 Perceptual and response model parameter results. *(A-E) No significant effects of COMT genotype on participants' perceptual belief updating, or on the sensitivity of participants' motor responses to their beliefs, were identified. Data are mean ± SEM.*

5.4.3.3 No identifiable effects of COMT genotype on response modulation

The response model output revealed no significant effects of the three *COMT* genotypes on participants' capacity to modulate their motor responses according to their perceptual estimates of sensory PE (β_1: $F_{2,113}$=1.72, p=0.183; Figure 5.7D) or phasic volatility (β_3: $F_{2,113}$=0.72, p=0.487; Figure 5.7E). The three *COMT* genotypes also had no significant impact on participants' general log(RTs) (β_0: $F_{2,113}$=0.57, p=0.568; Figure 5.7C). Similarly, comparing these response model parameters in Val/Val homozygotes and in individuals with higher DA neurotransmission (Met carriers) also indicated no effect of *COMT* genotype on the sensitivity of participants' motor responses to their sensory PE (β_1: t_{114}=1.57, p=0.118) or phasic volatility estimates (β_3: t_{114}=-1.17, p=0.244), or on general log(RTs) (β_0: t_{114}=1.06, p=0.292). Again, Levene's test confirmed equality of variances across groups for β_1 (F=1.12, p=0.293), β_3 (F=2.25, p=0.137) and β_0 (F=2.41, p=0.123).

For completeness, any effects of the three *COMT* genotypes on the sensitivity of participants' motor responses to their precision-weighted contingency PE were probed. No significant effects were identified (β_2: $F_{2,113}$=1.88, p=0.157). There was also no significant difference in the degree of post-error slowing demonstrated by individuals with each *COMT* genotype (β_4: $F_{2,113}$=2.89, p=0.060).

5.4.4 Control analyses

5.4.4.1 Model parameter correlations

Aside from two exceptions, Bayesian parameter averages (BPAs) for the different model parameters were only moderately correlated across groups (all absolute r≤0.352; Figure 5.8). As in Chapter 4, a higher correlation existed between the BPAs for β_0 (log(RT) constant) and $\beta_3(\mu_3)$ (the sensitivity of log(RTs) to phasic volatility estimates): r=-0.885. This negative correlation indicates that both the constant component of log(RT) and phasic volatility estimates had a similar effect on log(RTs). As mentioned in Chapter 4, this reflects the fact that, while including μ_3 as a predictor of log(RT) significantly improves model evidence, it is much less variable than the other predictors because volatility inevitably changes at a slower timescale than stimulus contingencies. In addition, a higher correlation existed between the BPAs for $\beta_1(\delta_1)$ (sensory PE) and $\beta_2(\varepsilon_3)$ (precision-weighted contingency PE): r=-0.7139. Again this reflects the similar effect the two parameters had on log(RTs).

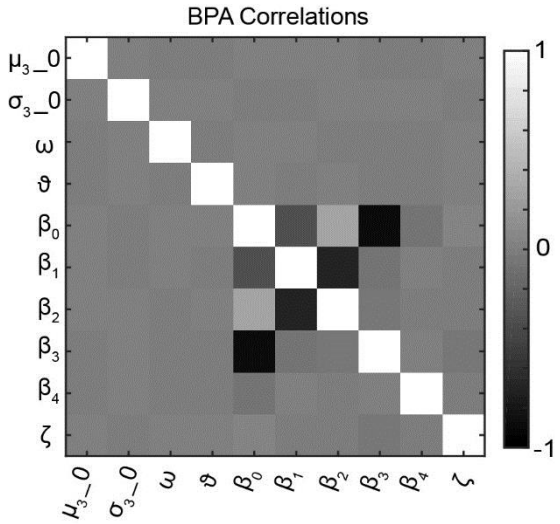

Figure 5.8 Model parameter correlations for Bayesian parameter averages (BPAs). *Note that μ_3_0 and σ_3_0 are the initial values of μ_3 (the phasic volatility estimate) and σ_3 (the uncertainty about the phasic volatility estimate) respectively.*

5.4.4.2 Sample size analysis

Since the effects of different genotypes on behavioural parameters are typically small, we had originally planned to recruit at least 400 participants to this experiment, in line with previous work (den Ouden et al., 2013). However, due to inconsistent data reporting by the genetics laboratory that assessed the DNA samples, testing had to be stopped prematurely. Nonetheless, for exploratory purposes, I used the data collected from the 116 tested participants to calculate the sample size (online materials: http://powerandsamplesize.com/Calculators/Compare-2-Means/2-Sample-Equality) that would be necessary to determine whether the sensitivity of motor responses to sensory PE (β_1) and to phasic volatility estimates (β_3) was altered by *COMT* genotype (Table 5.5). For simplicity, I computed the sample sizes that would be necessary to observe a significant effect of increased DA neurotransmission on these two response model parameters by comparing Met carriers to Val/Val homozygotes. A total of 374 participants would be required to observe an effect of *COMT* genotype on β_1 and 669 participants would be required to observe an effect on β_3, assuming a type I error rate (α) of 0.05 and a power (1-β, where β = type II error rate) of 0.8.

Parameter of interest	β_1 (sensory PE)	β_3 (precision-weighted contingency PE)
Observed mean (Val/Val); n=79	0.2162	-0.0726
Observed mean (Met carriers); n=37	0.1560	0.0637
Observed total standard deviation	0.1929	0.5856
Sampling ratio	2.14	2.14
Required sample size (Val/Val)	255	456
Required sample size (Met carriers)	119	213
Required total sample size	***374***	***669***

Table 5.5 Sample sizes required to observe an effect of COMT genotype on behaviour. *The calculations assume a comparison between two means (Group 1: Val/Val homozygotes; Group 2 = Met carriers). Note that type I (α) error rate is assumed to be 0.05 and power (1-β, where β = type II error rate) is set to 0.8.*

5.5 Discussion

Investigating learning and action during the PSRTT in a naïve cohort of 116 healthy human participants meant that the replicability of the results observed in Chapter 4's Placebo cohort could be assessed. Further, by using the novel instantiation of the HGF to probe individual perceptual belief updating and response modulation as a function of *COMT* genotype, it was possible to examine any effects of natural inter-individual variations in cortical DA neurotransmission on learning and action under uncertainty.

5.5.1 Replication of learning and response modulation under irreducible, estimation and volatility uncertainty

The naïve cohort of 116 healthy human participants who undertook the PSRTT in the present experiment showed learning and behaviour comparable to that displayed by the Placebo participants in Chapter 4's pharmacological study. Specifically, faster log(RTs) on trials with a high transition probability, which became even faster over the course of a contextual block, suggest that participants learned to predict the stimulus transitions, and to prepare appropriate motor responses, in the same way as the Placebo participants had in Chapter 4. In addition, findings of increased log(RTs) on post-error and post-infrequent trials were also replicated in the novel participant cohort.

As in Chapter 4, the HGF's perceptual model was found to track the true transition contingencies occurring in the PSRTT closely. Moreover, the same response model,

which captures trial-wise log(RT) as a function of a constant component of log(RT), sensory PE, precision-weighted contingency PE, phasic volatility estimate and post-error slowing, was found to be superior to alternatives. Together, these findings suggest that the PSRTT reliably evokes particular learning and behavioural processes in healthy individuals, which can be reliably captured by the novel instantiation of the HGF developed in Chapter 3 and applied in Chapter 4.

5.5.2 No identified effects of *COMT* genotype on perceptual belief updating or response modulation

In the present study, there was no evidence to suggest that *COMT* genotype, and hence the associated natural variations in cortical DA neurotransmission, had any effect on perceptual belief updating or motor response modulation between individuals. Indeed, no effects on learning and action were identified when the HGF's perceptual and response model parameters were compared across individuals with a Val/Val, Val/Met and Met/Met genotype, or when Val/Val homozygotes were compared to Met carriers.

5.5.3 Interpretation is limited by an insufficient sample size

However, interpretation of the behavioural genetics results in the present study is limited by an insufficient sample size. Genotypic effects on behavioural parameters are typically small, meaning that large sample sizes are commonly required to detect a behavioural effect (den Ouden et al., 2013). For this reason, I had originally planned to recruit at least 400 healthy participants. An additional complication is that distributions of different genotypes tend to be uneven across populations, as was indeed observed for the *COMT* genotype in the present experiment. This has the effect of reducing the statistical power to detect a behavioural effect as a function of genotype, further increasing the required sample size. Data collected from the 116 tested participants could be used to calculate the sample sizes needed to determine whether *COMT* genotype is associated with altered motor response sensitivity to sensory PE and phasic volatility estimates (the samples being 374 and 669, respectively). Therefore, while there is no evidence to suggest that *COMT* genotype modulated learning or behaviour in the present dataset, the experiment is not sufficiently powered to rule out the possibility that the $Val^{158}Met$ polymorphism does influence response modulation in uncertain environments.

5.5.4 Limitations of a behavioural genetics approach

While the behavioural genetics approach applied in the present experiment does hold the potential to identify inter-individual variations in DA-dependent learning and response modulation in a larger cohort of healthy individuals, there are several methodological

limitations. First, since the COMT enzyme actually degrades both DA and NA, any observed effect of *COMT* genotype on learning or action would have to be interpreted with caution. However, COMT's effect on NA levels is thought to be minor, at least in the PFC (Tunbridge et al., 2006), meaning that any confounding noradrenergic influences are likely to be minor. Indeed, *COMT* knockout mice show increased baseline frontal DA levels, with no effect on NA (Gogos et al., 1998). Moreover, administration of the COMT inhibitor tolcapone increases extracellular DA, but has no impact on NA, in the rat medial PFC following induction of neuronal catecholamine release (Tunbridge et al., 2004).

Second, categorising individuals according to genotype can often lead to unbalanced distributions of gender and ethnicity, as occurred in the present experiment. This means that confounding effects arising from distributional biases in genetic polymorphisms are possible. Indeed, an effect of gender on COMT activity has been identified in animal models (Gogos et al., 1998). Although statistical methods that account for the systematic differences that arise between experimental populations can help to address this issue, caution must be taken when basing conclusions on genetic polymorphisms that are rare and/or known to vary widely in frequency across males and females, or across different ethnic groups (Montana and Pritchard, 2004).

Third, the function and regulation of the neuromodulatory systems is intricate. A multitude of proteins, from receptors to transporters to degradative enzymes, mediate neurotransmission. As such, the individual genes that encode these proteins function in complex networks, and polymorphisms in each of them can influence neuromodulatory function. Therefore, while the Val^{158}Met polymorphism in the *COMT* gene might predict DA-mediated response modulation in volatile environments given a sufficient sample size, investigating the relative contributions of additional dopaminergic, noradrenergic and cholinergic genes could offer further insights. Indeed, there is evidence to suggest that the functional effects of the *COMT* polymorphism can be modulated by other dopaminergic polymorphisms, such as those in the *DAT1* and *DRD2* genes (Meyer-Lindenberg et al., 2006; Nackley et al., 2006; Diaz-Asper et al., 2008).

5.5.5 Future investigations of genotypic effects on learning and action under uncertainty

5.5.5.1 Alternative genetic targets

As discussed previously, the original aim of the present experiment was to assess the impact of *COMT*, *DAT1*, *DRD2*, *NET* and *ACHE* genotypes on perceptual belief updating and response modulation in uncertain environments. While it was only possible to probe

the effects of *COMT* here, the remaining genes remain sensible targets for future investigations. For instance, a VNTR in the *DAT1* gene, linked to altered expression of the DAT and thus altered striatal DA neurotransmission (Caron, 1996; Heinz et al., 2000; Lewis et al., 2001; Mill et al., 2002; van Dyck et al., 2005; van de Giessen et al., 2009; Frank and Fossella, 2011; Spencer et al., 2013), could offer a means by which to garner further insight into the contributions of DA to learning and action under uncertainty. In particular, contrasting any impact of *COMT* and *DAT1* genotypes on learning and response modulation during the PSRTT might enable cortical DA-mediated behavioural switching under volatility to be distinguished from striatal DA-mediated motoric responses to sensory PE. This would be of interest given the findings in Chapter 4 that pharmacological DA antagonism decreased the sensitivity of individuals' motor responses to their phasic volatility estimates, possibly due to impaired behavioural switching, but had no impact on responses to sensory PE. The latter result was surprising given previous literature demonstrating impaired reactions to unexpected events occurring in predictable contexts following both pharmacological and pathological DA depletions (Galea et al., 2012; Bestmann et al., 2014). Analysing behaviour in the PSRTT as a function of *DAT1* genotype would offer an alternative methodology with which to probe any DA-mediated motor responses to sensory PE.

Behavioural switching also varies with levels of striatal DA D2-receptor expression (Stelzel et al., 2010, 2013; van Holstein et al., 2011). A SNP in the *DRD2* gene is known to modify D2-receptor expression, particularly in the striatum (Noble, 2003). As such, investigating learning and action in the PSRTT as a function of *DRD2* genotype might offer still more intricate insight into the underlying dopaminergic processes. Given the proposed contributions of NA and ACh to learning under volatility and estimation uncertainty, genes that regulate noradrenergic and cholinergic neurotransmission also make sensible targets for future investigations. Indeed, while it is currently unclear exactly how two known polymorphisms in the *NET* and *ACHE* genes modulate neuromodulatory function, they may offer a means by which to speculatively study inter-individual differences in NA- and ACh-mediated learning, respectively.

5.5.5.2 Gene scoring

As mentioned above, the complexities of the neuromodulatory systems mean that a multitude of genetic polymorphisms can influence their function. For example, an individual might be a Met carrier for the *COMT* gene, a 10R carrier for the *DAT1* gene and an A2 carrier for the *DRD2* gene. As such, they might be expected to display relatively high cortical DA neurotransmission, low striatal DA neurotransmission and

increased DA D2-receptor expression. Further, it is possible that the relative levels of DA neurotransmission in particular brain regions might influence DA signalling in other brain regions, in addition to the function of the NA and ACh systems. This highlights the importance of considering the relative contributions of the various genes that modify neuromodulatory function when characterising the relative contributions of DA, NA and ACh to learning and action under uncertainty. Nevertheless, researchers must consider the statistical constraint of correcting for multiple comparisons when designing studies that aim to address the behavioural impact of multiple genotypes.

One possible way to address this is to adopt a gene scoring approach. Indeed, in their investigation of genetic variations in DA-mediated motor learning, Pearson-Fuhrhop et al. calculated a gene score that represented the additive effects of several polymorphisms with established effects on DA neurotransmission (Pearson-Fuhrhop et al., 2013). Genotypes thought to increase DA neurotransmission added 1 to the score while genotypes that decrease transmission added 0. As such, a higher DA polygene score corresponded to higher DA neurotransmission. A similar approach could be applied to assess whether the net effect of *COMT*, *DAT1* and *DRD2* polymorphisms on DA neurotransmission is associated with altered learning and response modulation during the PSRTT.

It should be noted that subtle insights into the relative cortical and striatal processes supporting these functions would be lost using such a method, unless the gene score weighted the contributions of different genes appropriately. Moreover, each of these genes is likely to only subtly alter the net function of an entire neuromodulatory circuit, meaning that elucidating how various genetic factors interact with each other is inherently difficult. Nonetheless, hypothesis-driven investigations of the functional impact of target genes and genetic interactions is likely the best approach. Given the multitude of potential genes that could influence learning and behaviour, unconstrained exploratory analyses of genetic interactions would require prohibitively large sample sizes to counter multiple comparisons and type I errors (i.e., false positives) (Purcell et al., 2009; Shi et al., 2009; Stefansson et al., 2009).

5.5.5.3 Combining genetics, pharmacological and neuroimaging approaches

A behavioural genetics approach to studying neuromodulatory contributions to learning and response modulation provides an important extension to insights offered by pharmacological manipulations. Indeed, the precise effects of pharmacological manipulations of neuromodulatory function can be unclear. For example, evidence from the animal literature suggests that DA D2-receptor antagonists such as haloperidol, the

drug utilised in Chapter 4, may actually exert their effects primarily via presynaptic autoreceptors (Richfield et al., 1989; Starke et al., 1989; Grace, 1995; Schoemaker et al., 1997; Schmitz et al., 2003; Frank and Fossella, 2011). Suppression of DA autoreceptor-mediated inhibitory feedback may actually increase phasic DA release, particularly in the basal ganglia where D2-receptors are highly expressed (Moghaddam and Bunney, 1990; Wu et al., 2002; Garris et al., 2003; Chen et al., 2005). It seems unlikely that this was the principal effect of haloperidol in Chapter 4 since the manipulation caused a general increase in RTs, echoing bradykinesia observed due to nigrostriatal DA depletions in Parkinson's disease (Galea et al., 2012). Nonetheless, the intricacies of the neuromodulatory effects caused by pharmacological interventions, and the compensatory mechanisms they may trigger, highlight the need to combine insights from pharmacology with alternative methodologies, such as behavioural genetics.

The correlational nature of behavioural genetics studies is another reason to combine them with insights offered by other methodologies. Indeed, corroboratory evidence acquired by adopting complementary methodologies will help to inspire confidence in the proposed contributions of different neuromodulatory systems to particular functions. For instance, Diaconescu et al. recently combined fMRI and behavioural genetics to demonstrate that low-level sensory PEs activate the dopaminergic midbrain and that these activations are influenced by the $Val^{158}Met$ polymorphism in the *COMT* gene, offering further weight to the notion that DA is implicated in signalling sensory PEs (Diaconescu et al., 2017). Together, behavioural genetics, psychopharmacological and neuroimaging studies, both in healthy individuals and in patients with known neuromodulatory dysfunction, can help to elucidate the neuromodulatory contributions of DA, NA and ACh to learning and action in uncertain environments.

A further reason to combine behavioural genetics and psychopharmacology is that individual behavioural responses to pharmacological manipulations can depend strongly on baseline levels of DA, NA and ACh neurotransmission (Kimberg et al., 1997; Mehta et al., 2004a; Roesch-Ely et al., 2005; Frank and O'Reilly, 2006; Cools et al., 2007b; Clatworthy et al., 2009). For example, pharmacological DA D2-receptor stimulation generally improves task performance in individuals with low baseline working memory span (Kimberg et al., 1997; Frank and O'Reilly, 2006), high impulsivity (Cools et al., 2007b) or low baseline DA synthesis (Cools et al., 2009), but impairs performance in those showing the opposite baseline trait. Since multiple genes are thought to modulate baseline neuromodulatory function, there is strong reason to predict that individual differences in dopaminergic, noradrenergic and cholinergic drug effects are, at least in part, genetic. Indeed, *DRD2* genotype has been shown to predict the direction of an

individual's neural and behavioural responses to pharmacological DA D2-receptor stimulation (Cohen et al., 2007). I will return to this concept in Chapter 7.

5.5.6 Conclusion

In conclusion, by employing the PSRTT to interrogate learning and action in a naïve cohort of healthy participants exposed to dynamic, probabilistic environments, it was possible to replicate the behaviour displayed by the Placebo participants in Chapter 4. Moreover, the novel instantiation of the HGF showed a verifiable capacity to capture individual perceptual belief updating and response modulation within a unified computational framework of irreducible, estimation and volatility uncertainty. As such, the PSRTT and HGF appear to offer a robust means by which to investigate learning and response modulation in uncertain environments. While previous literature linking cortical DA neurotransmission and behavioural flexibility suggests that it is reasonable to predict that *COMT* genotype might modify the modulation of motor responses by phasic volatility estimates, future work will need to confirm whether this effect can be detected using the PSRTT in a larger participant cohort. Additional genetic, pharmacological and neuroimaging investigations of dopaminergic, noradrenergic and cholinergic neurotransmission will verify the generality of any behavioural effects to different behavioural tasks with and without learning, reward, prediction and action.

6 Dynamic noradrenergic computations of uncertainty

*This chapter was, in part, motivated by work presented in de Berker AO, Rutledge RB, Mathys C, **Marshall L**, Cross GF, Dolan RF & Bestmann S. (2016) Computations of uncertainty mediate acute stress responses in humans. Nature Communications. 7:10996.*

6.1 Abstract

Noradrenaline (NA) has been proposed to play an important role in learning under the uncertainty that arises from environmental volatility. Parallel lines of work have linked subjective uncertainty computations, and noradrenergic activity in the locus coeruleus (LC), to changes in pupil diameter. In this chapter, I combine a probabilistic learning task, pupillometry, pharmacological manipulations and the Hierarchical Gaussian Filter (HGF) model to characterise the impact of subjective beliefs and NA on pupillary dynamics in 90 healthy human participants. Baseline pupil diameter was found to reflect an individual's belief about the current relationship between environmental events. Dynamic pupillary dilation tracked both uncertainty and surprise arising from the probabilistic relationship between environmental events. Pharmacological manipulations of NA modulated pupillary responses to uncertainty and volatility estimates. Collectively, the results provide empirical support for the notion that pupil diameter offers an indirect measure of individual dynamic noradrenergic computations of uncertainty and volatility. Importantly, they also highlight the need for unified behavioural and computational frameworks in characterising the relative contributions of subjective beliefs and neuromodulatory dynamics to pupil dilation.

6.2 Introduction

For several decades, pupil dilation at constant luminance has been considered a marker of central arousal (Hess and Polt, 1964; Kahneman and Beatty, 1966; Bradshaw, 1967; Kahneman et al., 1967; Beatty, 1982). More recently, it has been proposed that pupil diameter might offer an indirect measure of noradrenergic neuronal activity in the LC (Rajkowski et al., 1993; Phillips et al., 2000b; Aston-Jones and Cohen, 2005a; Murphy et al., 2014; Varazzani et al., 2015; Joshi et al., 2016). Further, as we saw in Chapter 4, NA has been linked to learning under volatility uncertainty (Yu and Dayan, 2005; Payzan-LeNestour et al., 2013; Marshall et al., 2016). Inspired by these findings, researchers have started to probe whether transient changes in pupil diameter can be used as a proxy for physiological autonomic processes that occur during behavioural tasks, including those requiring NA-mediated learning under environmental uncertainty (Siegle et al., 2003; Aston-Jones and Cohen, 2005a; Critchley, 2005; Satterthwaite et al., 2007; Einhäuser et al., 2008; Hupé et al., 2009; Einhauser et al., 2010; Gilzenrat et al., 2010; Privitera et al., 2010; Jepma and Nieuwenhuis, 2011; Preuschoff et al., 2011; Fiedler and Glöckner, 2012; Nassar et al., 2012; Wierda et al., 2012; Eldar et al., 2013; de Gee et al., 2014; Browning et al., 2015; de Berker et al., 2016; Korn et al., 2016; van den Brink et al., 2016; Urai et al., 2017). The sensitivity of the pupil to such processes means that pupillometry might offer a simple, non-invasive and cost-effective tool with which to measure individual noradrenergic computations of uncertainty, without the need for pharmacological interventions or behavioural genetics analyses.

6.2.1 Pupil diameter as an indirect measure of noradrenergic neurotransmission

As discussed in detail in Chapter 1, converging bodies of electrophysiological (Rajkowski et al., 1993; Aston-Jones and Cohen, 2005a; Varazzani et al., 2015; Joshi et al., 2016), pharmacological (Phillips et al., 2000c) and human neuroimaging (Samuels and Szabadi, 2008; Murphy et al., 2014) evidence suggest a relationship between NA and pupil dilation under constant luminance.

6.2.2 A proposed link between pupil diameter and perceptual beliefs

A parallel line of work has sought to establish whether pupil diameter reflects an individual's perceptual estimates by integrating pupillometry into studies of human learning and behaviour (Gilzenrat et al., 2010; Jepma and Nieuwenhuis, 2011). In particular, during the last five years, researchers have focused on developing quantitative models to formally test the hypothesised association between human pupil

dilation and the perceptual estimates underlying learning and behaviour in uncertain environments. Indeed, a range of studies has investigated the modulation of pupil diameter by perceptual quantities such as uncertainty, prediction error (which the pupil literature commonly conceptualises as surprise), and volatility (Preuschoff et al., 2011; Nassar et al., 2012; de Gee et al., 2014; Browning et al., 2015; de Berker et al., 2016). The results of these studies have offered varying evidence to suggest that pupil dilation is associated with each of these perceptual estimates.

6.2.3 Pupil diameter as a proxy for dynamic noradrenergic uncertainty computations

Given my finding in Chapter 4 that NA modulates learning under the uncertainty arising from environmental volatility, the notion that pupil dilation can be used as a proxy for dynamic noradrenergic uncertainty computations is appealing. However, the previous investigations of pupillary responses to perceptual estimates have been heterogeneous: they used different behavioural paradigms that exposed participants to different forms of environmental uncertainty, and the investigators probed the impact of different combinations of perceptual beliefs on pupil diameter. As such, it is difficult to isolate the contribution of particular perceptual estimates to pupil diameter with confidence. Therefore, in this chapter, I combine a probabilistic learning task, pupillometry and the HGF model to assess the impact of irreducible uncertainty, surprise and volatility on pupil diameter. Further, by utilising two pharmacological manipulations of NA, I causally assess whether any pupillary responses to these perceptual beliefs are under dynamic noradrenergic modulation.

6.3 Methods

6.3.1 Participants

90 healthy participants (39 male, aged 19-38 years, 83 right-handed) with normal hearing and normal or corrected-to-normal vision took part in this study after giving written informed consent. The experimental protocol was approved by the UCL Research Ethics Committee. The following exclusion criteria applied: history of neurological or psychiatric disorder, baseline blood pressure below 100/60, intake of medication (other than contraceptives), self-reported regular smoking, self-reported recreational drug use, and current participation in other pharmacological studies. Following a screening interview to rule out intolerances or contraindications, the study clinician assigned participants pseudorandomly (i.e., ensuring a balanced distribution of gender, age and body weight)

to receive a NA antagonist, a NA reuptake inhibitor or a placebo. The experimenter (L.M.) was blind to the drug conditions.

6.3.2 General procedure

A double-blind, between-subjects design was employed. Each participant attended one experimental session during which they received a single, oral dose of one of the following: 4mg reboxetine (selective NA reuptake inhibitor; NA+ group), 1mg prazosin (α1-adrenoceptor antagonist; NA- group), or a placebo. Doses were selected in line with previous studies showing clear behavioural and neurophysiological effects (Dostert et al., 1997; Meintzschel and Ziemann, 2006; de Martino et al., 2007; Jepma et al., 2010; Korchounov and Ziemann, 2011; Marshall et al., 2016). On arrival, participants completed computerised versions of the Digit Span test, Barratt Impulsiveness Scale (BIS-11) (Patton et al., 1995), Domain-Specific Risk-Taking (DOSPERT) Scale (Blais and Weber, 2006) and Cognitive Failures Questionnaire (CFQ) (Broadbent et al., 1982). Participants also self-reported their baseline mood (alertness, calmness and contentedness) with 16 visual analogue scales (VAS) (Bond and Lader, 1974), and had their baseline heart rate (HR) and blood pressure (BP) measured. To assess any subjective and/or physiological drug effects, the VAS, HR and BP measurements were repeated before participants started the behavioural task and again once they completed it.

All participants were administered two tablets thirty minutes apart. Two different active drug administration times were used, based on previous pharmacokinetic data, so that participants undertook the behavioural task when the drug was at its most active (Dostert et al., 1997; de Martino et al., 2007; Jepma et al., 2010). Reboxetine was administered two hours before the main behavioural task (Time A; Figure 6.1A), and prazosin 1.5 hours in advance (Time B). Participants in the NA+ (reboxetine) group received a placebo at Time B, while participants in the NA- (prazosin) group received a placebo at Time A. Participants in the Placebo group received a placebo tablet at Time A and at Time B. Participants were told that they would receive either two placebo tablets, or one active drug and one placebo tablet. An independent clinician administered the drug or placebo while the experimenter was away from the testing room. For comparable drug absorption rates, participants were asked not to eat for at least one hour before Time A.

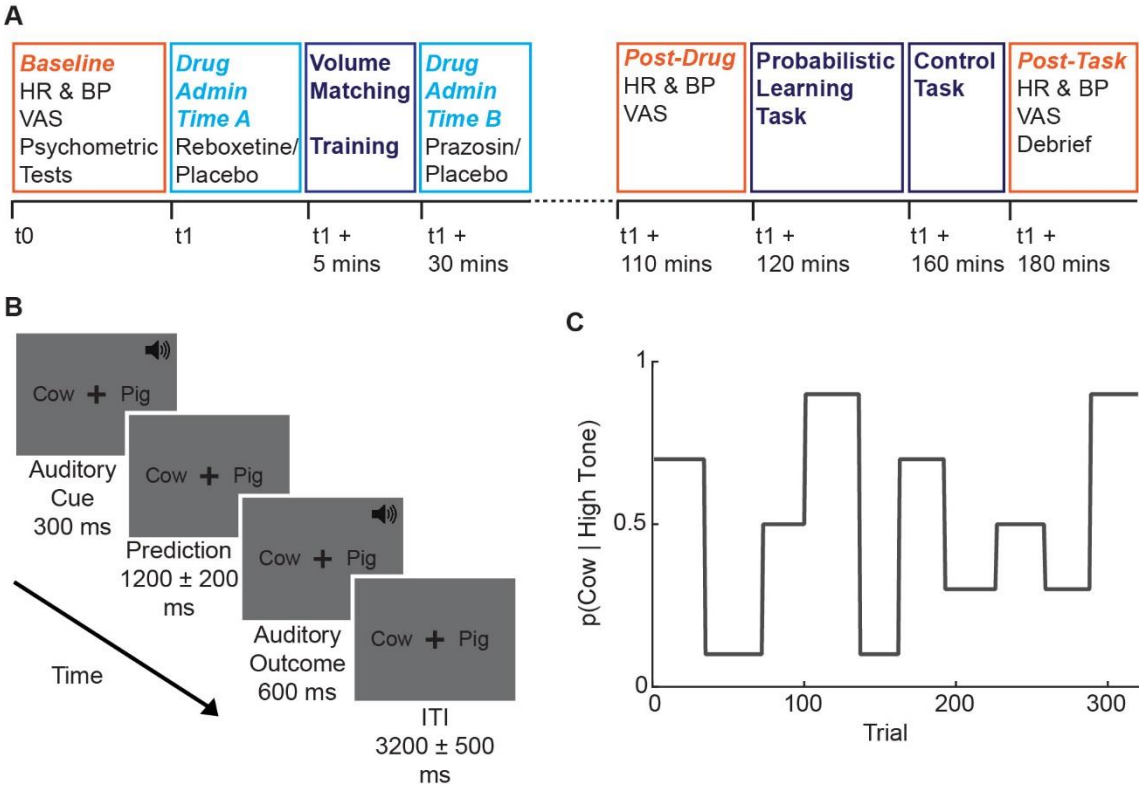

Figure 6.1 Task design. *(A) Timeline for each experimental session. At baseline, heart rate (HR) and blood pressure (BP) measurements were taken, and participants self-reported their alertness, contentedness and calmness via visual analogue scales (VAS) (Bond and Lader, 1974), and undertook a battery of psychometric tests to assess working memory, impulsivity, risk-taking and distractibility. HR, BP and VAS measures were repeated before and after completing the behavioural task. Due to different times-to-peak plasma concentration across drugs, two different drug administration times (Time A and Time B) were used so that participants undertook the behavioural task when the drugs were at their most active. Participants in the active drug-groups received a placebo tablet at the administration time at which they did not receive an active drug. Placebo participants received two placebo tablets: one at Time A and one at Time B. (B) Trial sequence. Throughout the task, an isoluminant grey screen was displayed with a central black fixation cross and the words "Cow" and "Pig" on each side. On each trial, an auditory low-pitch (450Hz) or high-pitch (1000Hz) cue was presented for 300ms. Participants were required to make a speeded button-press response within a 1200 ± 200ms window to indicate their prediction about which auditory outcome would follow. The auditory outcome was either the word "cow" or the word "pig", presented for 600ms. This was followed by an intertrial interval (ITI) lasting 3200 ± 500ms. (C) The probabilities governing cue:outcome relationships shifted unpredictably over time, introducing volatility and thus producing fluctuations in uncertainty.*

6.3.3 Probabilistic learning task

Participants sat in a darkened room facing an isoluminant computer screen. They were instructed to rest their left and right index fingers on two response buttons, and to maintain this position throughout the task. Participants were asked to maintain fixation on a black fixation cross presented in the centre of a grey screen at a viewing distance of 60cm. During the probabilistic learning task (PLT), the diameter of the left pupil was measured using an infrared ASL Eye-Trac 6 System (Applied Science Laboratories, USA), sampled at 120Hz. To minimise movement, participants sat with their head supported by a forehead- and chin-rest (Figure 2.5).

The PLT was closely modelled on a task used in three recent studies (den Ouden et al., 2010; Iglesias et al., 2013; de Berker et al., 2016), but used auditory rather than visual stimuli so as to eliminate any effects of luminance changes on pupil diameter. Each participant completed a set of 320 trials. On each trial, participants were presented, via stereo headphones, with one of two auditory cues: a low-pitch (450Hz) or high-pitch (1000Hz) tone. The cue was presented for 300ms before participants were asked to make a prediction, signalled with a speeded button-press response, as to which auditory outcome (the word “cow” or the word “pig”) would follow (Figure 6.1B). This decision was made under time pressure, with a timeout period averaging 1200ms (± 200ms). The auditory outcome was followed by an intertrial interval (ITI) averaging 3200ms (± 500ms). The durations of the decision period and the ITI were jittered on each trial so that the pupil responses could be maximally divorced from different events. Jitter was implemented using a uniform distribution, discretised into chunks of a size determined according to the size of the interval in question.

The probabilistic mapping between cue and outcome shifted over the course of the experiment (Figure 6.1C), requiring participants to constantly track the cue:outcome relationship over time in order to maximise their proportion of correct predictions. This resulted in fluctuations in the level of uncertainty about the outcome, given the cue. Each session of 320 trials was divided into 10 blocks of different cue:outcome probabilities, and of lengths that varied between 26 and 38 trials. The transitions between these blocks were not made explicit to the participant. The probabilities governing each block varied from highly biased (0.9/0.1), through moderately biased (0.7/0.3) to unbiased (0.5/0.5), allowing the effect of predictability (Iglesias et al., 2013) on pupil diameter (de Berker et al., 2016) to be examined. Each of the biased probability blocks was repeated four times (two for each bias direction, i.e. 0.7/0.3 and 0.3/0.7) and the unbiased blocks were repeated twice.

Participants were told that the cue:outcome probabilities would shift unpredictably over time and might sometimes be completely random, but they were uninformed as to the frequency of switches. The task therefore required participants to track the same three forms of uncertainty we have encountered in previous chapters: *irreducible uncertainty* arising from the inherent randomness of the probabilistic relationships between cues and outcomes, *estimation uncertainty* arising from incomplete knowledge of those probabilistic relationships within each block, and *volatility uncertainty* maintained by the unsignalled instability of the relationships over time (Behrens et al., 2007; Mathys et al., 2011, 2014; Payzan-LeNestour and Bossaerts, 2011; Iglesias et al., 2013; Payzan-LeNestour et al., 2013; Hauser et al., 2014; Vossel et al., 2014a, 2014b, 2015; de Berker et al., 2016; Diaconescu et al., 2017).

The visual display remained stable throughout, with the word “cow” displayed to the left of the fixation cross and the word “pig” to the right. Rest periods lasting 90 seconds occurred every 65 trials, orthogonal to cue:outcome probability switches. Before starting the PLT, participants underwent a volume matching procedure to ensure that they perceived the different auditory cues and outcomes as equally loud, and a training block to familiarise themselves with making button-press predictions in response to auditory cues (see sections 6.3.5 and 2.1.2.3 for details).

To encourage task engagement, participants were paid a base rate of £25 and informed that they would receive an extra £5 if they could make correct predictions on more than 68% of trials. This threshold was based on the average number of correct predictions made by participants who undertook a similar experiment conducted in my work with de Berker et al. (de Berker et al., 2016). It was not explicitly signalled to the participant whether each outcome reflected a correct or incorrect prediction on their part.

6.3.4 Control task

After completing the PLT, participants undertook an additional control task (CT) consisting of two blocks of 30 trials each. The CT was identical to the PLT except that this time there was no uncertainty about which auditory outcome would follow which auditory cue. Rather, participants were explicitly told at the beginning of each block which cue and which outcome would occur on each of the next 30 trials. One cue was paired with one outcome for the first block, and the other cue was paired with the other outcome for the second block. On hearing the auditory cue at the start of each trial, participants were required to make a speeded button-press response to indicate their “prediction” as to which outcome they knew would follow. Timings for the CT were identical to those used in the probabilistic task. The exact pairings between cues and outcomes, and their

assignment to the first or second block of the CT were counterbalanced across participants. As before, the diameter of the left pupil was recorded online at 120Hz. As one would expect, participants were highly accurate during the CT: the mean percentage of correct “predictions” was 100% for participants in the NA- and NA+ groups, and 98.3% for those in the Placebo group.

6.3.5 Volume matching

Before undertaking the behavioural task, participants underwent an adaptive, two-alternative forced choice procedure in order to match the subjective loudness of the two cues and the two outcomes used in the PLT. On each trial, participants were played the two cues or the two outcomes in succession and asked to report whether the second sound was louder or quieter than the first. The volumes of the low-pitch cue and the outcome “cow” were kept constant throughout. The level of the high-pitch cue and the outcome “pig” were varied according to a maximum likelihood procedure (Green, 1993; Soranzo and Grassi, 2014) to obtain estimates of the attenuation required for subjective volume equality. For the two cues and the two outcomes, three adaptive runs of 10 trials were performed. The average final attenuation values for the cues and outcomes were used in the PLT and CT.

6.3.6 Training

Participants were trained on the PLT before starting it. During four training blocks of five trials each, participants familiarised themselves with making predictions by button press following the presentation of auditory cues. Participants were told at the start of each training block which cue and which outcome would be presented on each of the following five trials. After the outcome was presented, they were provided with visual error feedback. Each combination of cue and outcome was presented across the four training blocks. The order of the four training blocks (i.e., the pairings between each cue and each outcome) was counterbalanced across participants. To familiarise themselves with the timings of the PLT, participants then completed 12 practice trials without error feedback. On each trial there was a 50% probability that either cue would be followed by either outcome.

6.3.7 Pupillometry

The ASL Eye-Trac system calculates pupillary gaze by measuring the distance between the location of a participant’s pupil and corneal reflection (CR). For each participant, the eyetracker was calibrated to account for inter-participant differences in the relationship between the pupil and CR. Each participant was instructed to sequentially fixate nine

calibration points arranged in a 3x3 square on the computer screen ahead of them. The central calibration point was positioned at the location of the centre-point of the fixation cross used during the PLT, which was horizontally centred on the computer screen and in line with the participant's line of vision when looking straight ahead. During the PLT and CT, calibration was repeated after each rest period to adjust for any subtle differences in head position. In order to align the pupil diameter timecourse with experimental events occurring in the PLT and CT (i.e., the precise timing of cue, response and outcome onsets), triggers were sent via the testing computer's parallel port to the eyetracker system. Pupil diameter was sampled at 120Hz.

6.3.8 Model-agnostic analyses

A series of conventional, model-agnostic analyses of behaviour were first conducted to assess whether participants showed evidence of learning the underlying cue:outcome relationships during the PLT, and whether learning was influenced by the pharmacological interventions.

6.3.8.1 Accuracy and decision time

Accuracy was defined as the percentage of correct responses. Decision time was calculated as the time between cue onset and the subsequent button-press response made to indicate the predicted outcome.

First, accuracy and decision times during the unbiased (0.5/0.5) blocks were assessed to verify that they were equivalent across drug-groups. Here no probabilistic advantage arose from the cue:outcome relationship, meaning that any participant who fully understood the task requirements and was actively engaged in the PLT would be expected to perform at close to chance level. Indeed, any drug effect would indicate a non-specific effect on behaviour rather than an effect specific to altered uncertainty computations.

To obtain a model-agnostic indication of learning across the course of the probabilistic blocks, a median split was performed on each block. A 3 bias (high/moderate/none) x 2 time (Early/Late) x 3 drug repeated-measures analysis of variance (RM-ANOVA) was used to compare accuracy on Early (first half of each block) and Late (second half of each block) trials at each bias level, and between drug-groups. As two bias directions had been used for each bias level, I collapsed across highly biased blocks (0.9/0.1 and 0.1/0.9), and across moderately biased blocks (0.7/0.3 and 0.3/0.7). A second bias x time x drug RM-ANOVA was applied to assess decision times in the same way.

6.3.8.2 Performance score

Since improved performance on the PLT is reflected by a higher accuracy rate and lower decision times (Volkmann, 1934; Yeung et al., 2004; Fetsch et al., 2014), a performance score that captured these two components was calculated as follows:

$$Performance\ Score = \frac{Accuracy}{Decision\ Time}$$

A higher performance score reflects improved performance owing to an increased accuracy and/or decreased decision time. RM-ANOVAs were applied to compare mean performance scores across the three bias levels, Early and Late Trials, and across drug-groups.

6.3.9 Model-based analyses

A three-level instantiation of the HGF model was applied to the behavioural data to quantify participants' (approximate) inferences and subjective expectations about trial-wise auditory stimulus outcomes.

To recap Chapter 3, the original instantiation of the HGF applied here consists of a perceptual model (i.e., a generative and recognition model) that tracks an individual's learning of the PLT's structure: the trial-wise auditory stimulus outcomes at level 1, the probabilistic relationship between cue and outcome at level 2, and the volatility of the cue:outcome relationship at level 3. Trial-wise trajectories of a participant's estimates at each level evolve according to the predictions made and outcomes experienced by that individual (Figure 6.2). At levels 2 and 3, these beliefs are represented as Gaussian distributions characterised by a mean (μ) and a variance (σ). This framework naturally captures *irreducible uncertainty* resulting from the probabilistic relationships between cues and outcomes, *estimation uncertainty* resulting from imperfect knowledge of these probabilistic relationships, and *volatility uncertainty* reflecting the instability of these relationships over time.

Importantly, the model does not assume fixed learning across participants but rather contains participant-specific parameters that couple the hierarchical levels and allow for individual expression of approximate Bayes-optimal learning. ϑ determines the speed of learning about phasic volatility, and ω captures how rapidly individuals update their beliefs about the cue:outcome relationship.

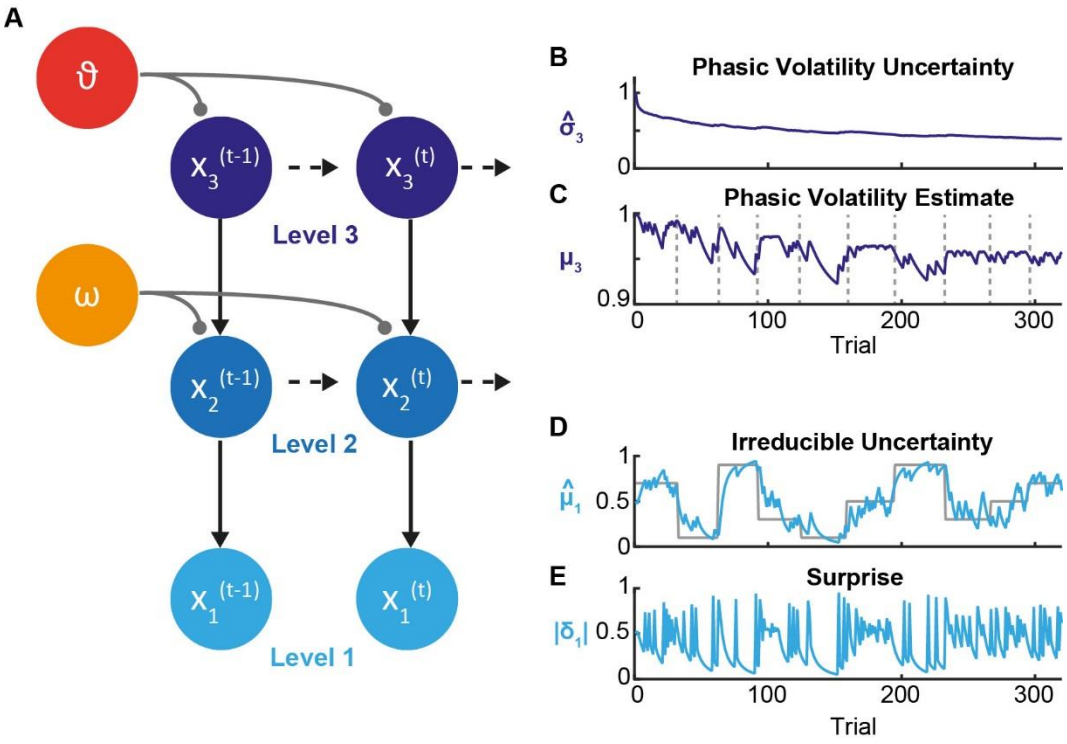

Figure 6.2 The Hierarchical Gaussian Filter (HGF). *(A) The model tracks an individual's learning of the task's structure across three levels. State x_1 represents trial-wise auditory stimulus outcomes, x_2 the probabilistic relationship between cue and outcome, and x_3 the phasic volatility of this relationship, where t is the current trial number. Participants hold and update beliefs about the true quantities at each level. (B-E) Examples of the trial-wise dynamics at levels 1 and 3 for Placebo Participant 2. At level 1, irreducible uncertainty ($\hat{\mu}_1$) results from a sigmoid transformation of the estimated probabilities (μ_2) represented at level 2. As such, $\mu_1^{(t)}$ reflects the participant's current belief about the true cue:outcome probabilities (x_2; grey line in D), which the participant tracks closely. Irreducible uncertainty gives rise to trial-wise estimates of surprise, which is mathematically equivalent to sensory prediction error ($|\delta_1|$; E). At level 3, μ_3 reflects the participant's belief about the true phasic volatility (x_3). μ_3 tends to increase following the true switches in cue:outcome probability (marked by grey dashed lines in C), and decreases over the course of the highly biased blocks as the participant learns the new cue:outcome relationship and thus perceives the environment as increasingly stable. An individual's uncertainty about their predicted phasic volatility estimate is captured by $\hat{\sigma}_3$ (B).*

6.3.9.1 Parameters of interest

To probe dynamic noradrenergic responses to uncertainty, four key parameters from the HGF were assessed. In my previous work with de Berker et al., we found evidence to

suggest that decision times are modulated by low-level irreducible uncertainty and that pupil diameter appears to track both irreducible uncertainty and surprise (de Berker et al., 2016). Therefore, the effect of irreducible uncertainty and surprise on behaviour and pupil diameter was also assessed in the current study. Irreducible uncertainty ($\hat{\mu}_1$) arises from the inherent randomness of the probabilistic cue:outcome relationships. It is a quantity closely related to entropy, with an inverted-U relationship to probability that peaks at p=0.5. Trial-wise values are equivalent to 1-$\hat{x}_1$, where $\hat{x}_1$ is the probability of the predicted outcome. Irreducible uncertainty gives rise to sensory prediction error ($|\delta_1|$), which the pupil literature commonly describes as surprise.

In Chapter 4, I found evidence to suggest that NA influences learning of uncertain events arising from unexpected changes in the environment, a finding in line with previous literature linking NA to uncertainty arising from environmental volatility (Yu and Dayan, 2005; Payzan-LeNestour et al., 2013; Marshall et al., 2016). In a parallel line of research, it has been proposed that pupil dilation reflects an individual's subjective volatility estimates (Browning et al., 2015). Moreover, pupil dilation has been suggested to offer an indirect measure of NA activity (Murphy et al., 2014; Varazzani et al., 2015; Joshi et al., 2016). Therefore, I also assessed the impact of two further parameters from level 3 of the HGF on pupil diameter. First, the trial-wise estimate of phasic volatility (μ_3), and second the uncertainty about phasic volatility beliefs ($\hat{\sigma}_3$).

Inspired by literature linking volatility estimates to learning rate (Behrens et al., 2007; den Ouden et al., 2010; Browning et al., 2015; Jepma et al., 2016), I also assessed learning rates (α_1) at level 1 across drug-groups.

6.3.9.2 Model fitting

The HGF model was implemented using the 'tapas_hgf_binary' code contained in the HGF Toolbox (http://www.translationalneuromodeling.org/tapas/). Where priors were required, they were generated by running a Bayes optimal version of the model (using the function 'tapas_bayes_optimal_binary_config'), under suitably uninformative priors. The resulting posterior estimates were then used to define the priors for the subsequent inversion of the full model given the behavioural data (see Table 6.1). In other words, the prior means in the empirical data analysis corresponded to those parameter values for which the stimulus sequence would generate minimal surprise (in an observer with the aforementioned uninformative priors). Note that these priors are in line with those used in our previous study using a visual version of the PLT (de Berker et al., 2016).

6.3.9.3 Decision times according to beliefs

To verify the HGF's capacity to capture participants' beliefs about the true cue:outcome probabilities, trials were binned according to 10 evenly-spaced irreducible uncertainty belief levels (model parameter $\hat{\mu}_1$) ranging from 0/1 to 1/0. Since two bias directions had been used in the PLT, the data were symmetrical. Therefore, I next collapsed across

Parameter	Notes	Prior	
ϑ	Metavolatility belief parameter; controls the step size of the Gaussian random walk at level 3. Estimated in logit space.	Mean Variance	0 16
ω	Tonic volatility belief parameter; a constant component of the learning rate at level 2.	Mean Variance	-3 16
Stimuli (x_1)	The stimulus predictions are a sigmoid transformation of the probabilities represented in x_2, and so do not have a starting value.	μ_1: Mean Variance	 NaN NaN
		σ_1: Mean Variance	 NaN NaN
Probabilities (x_2)	A starting value of 0 implies neutrality between outcomes.	μ_2: Mean Variance	 0 0
		σ_2: Mean Variance	 log(0.1) log(1)
Volatility (x_3)	The absolute starting value of x_3 is arbitrary as changes in fitted parameters affect scaling.	μ_3: Mean Variance	 1 0.1
		σ_3: Mean Variance	 log(1) 0

Table 6.1 A summary of HGF parameters and priors.

All priors are specified in the space in which they are estimated. For an account of how this relates to the native space of that parameter, please refer to Chapter 3 and to the original description of the model (Mathys et al., 2011).

equivalent beliefs in each bias direction to create five bias categories (i.e., 0-0.1/0.9-1, 0.1-0.2/0.8-0.9. 0.2-0.3/0.7-0.8, 0.3-0.4/0.6-0.7, 0.4-0.5/0.5-0.6). A 5 bias x 3 drug RM-ANOVA was used to compare mean decision times across these five belief levels and between drug-groups. Since the model is informed of participants' trial-wise predictions, but not their decision times, an increase in decision time as irreducible uncertainty, and hence the belief about cue:outcome probabilities, approach 0.5/0.5 would indicate that the HGF had captured participants' beliefs well, assuming participants showed typical

decision time slowing with increasing uncertainty (Volkmann, 1934; Yeung et al., 2004; Fetsch et al., 2014).

6.3.9.4 Learning rate

As in Chapter 4, I examined the learning rate (model parameter α_1), across the course of the PLT and at true context change-points, i.e., following a switch in cue:outcome probability. Where differences in learning rate at change-points were calculated, the mean α_1 for the last three trials of the previous context were subtracted from the mean α_1 for the first two trials of the next context. ANOVAs were applied to assess whether learning rates differed across drug-groups, and whether the noradrenergic manipulations modulated how learning rates changed at context change-points.

6.3.10 Statistical analyses of behavioural data

In reporting statistical differences, a significance threshold of α=0.05 was used. Where assumptions of sphericity were violated (Mauchly's test $p<0.05$), the Greenhouse-Geisser correction was applied. Since a significant time x drug interaction on self-reported alertness was identified (see section 4.4.5.1 for details), the participant-specific difference in alertness between baseline and the time corresponding to peak drug concentration, *Δalertness*, was used as a covariate in all analyses to control for any inter-participant variability in subjective drug effect.

For comparisons across the three drug-groups, partial eta-squared (η_p^2) is reported as the effect size. The key experimental question pertained how noradrenergic manipulations influence behaviour, and pupillary responses to uncertainty estimates, compared to placebo. Therefore, for behavioural data, planned comparisons were made between the two active drug-groups (NA- and NA+) and the Placebo group. Here a Benjamini-Hochberg correction for two pairwise comparisons was applied to account for the false discovery rate (FDR) (Benjamini and Hochberg, 1995). For pairwise comparisons, Cohen's *d* is reported as the effect size.

6.3.11 Analysis of pupil diameter

For the purpose of analysis, pupil data were exported using ASL software and then imported into Matlab (MathWorks, USA). Blinks (defined as pupil losses lasting ≤300ms) were detected using a custom-made algorithm and removed by linear interpolation of samples 50ms either side of the blink. Additional artefacts (<50ms) were identified in the data by taking the first derivative of the pupil series to detect rapid (sample-wise) changes in pupil diameter measuring >10% of the maximum pupil diameter. These

artefacts were removed by the same linear interpolation method. Pupil losses lasting longer than 300ms were removed by NaN padding. The interpolated pupil time series were low-pass filtered (4Hz, 3rd order Butterworth) (de Gee et al., 2014; de Berker et al., 2016), detrended, and z-scored. In line with Browning et al., 2015, trials in which more than 50% of the eyetracking data were interpolated or lost were not used in subsequent analyses (mean ± SEM: 2.8 ± 0.8% of trials for NA-, 1.4 ± 0.6% for Placebo, and 2.5 ± 1.1% for NA+).

6.3.11.1 Event-related analysis of pupil diameter

To gain insight into the role played by uncertainty, the pupillary response was epoched by trial (i.e., from -200ms from cue onset on trial *t* to -200ms from cue onset on trial *t+1*). Each epoch was baseline-corrected by subtracting the mean of all pre-trial pupil diameter values in the window from -200 to 0ms from cue onset from the pupil diameter trajectory for that trial. Each trial-wise, baseline-corrected pupil diameter trajectory was separated into three epochs: 1) -0.2 to 0.3s from cue onset, 2) -0.2 to 1s from response onset (i.e., the button-press that indicated a participant's decision), and 3) -0.2 to 3s from outcome onset. Any trials on which participants failed to indicate a decision by button-press before outcome presentation was excluded. The percentage of missed button-press responses was very low across drug-groups (mean ± SEM: 0.55 ± 0.17% of trials for NA-, 0.34 ± 0.09% for Placebo, and 0.76 ± 0.29% for NA+).

A one-way ANOVA was applied to assess peak pupil diameter during the post-response and post-outcome periods across drug-groups. For the post-response period, the peak pupil was calculated as the mean of a 100ms window spanning the time of the average peak pupil measurement. For the outcome period, a 250ms window was used. Trial-wise pupillary responses during the PLT were contrasted with trial-wise pupillary responses during the CT, which was identical in structure to the PLT but gave rise to no uncertainty about the cue:outcome relationship. Again, the percentage of (excluded) missed button-press responses was very low in the CT across drug-groups (mean ± SEM: 0.33 ± 0.04% of trials for NA-, 0.05 ± 0.02% for Placebo, and 0.10 ± 0.05% for NA+).

6.3.11.2 Pupil diameter at baseline

Within the framework of the HGF model, information about irreducible uncertainty on the current trial is available to participants before the trial begins since it is computed on the basis of trial history (Mathys et al., 2011). Therefore, in line with my previous work (de Berker et al., 2016), baseline pupil diameter, computed as the mean of all pre-trial values in the window from -200ms to 0ms from cue onset, was interrogated in order to determine

whether it reflected participants' current beliefs, and whether this relationship differed across drug-groups. As with the decision time data, trials were first binned according to 10 evenly-spaced belief levels (model parameter $\hat{\mu}_1$) ranging from 0/1 to 1/0. For statistical comparisons, I collapsed across equivalent probability beliefs in each bias direction to create five bias categories (i.e., 0-0.1/0.9-1, 0.1-0.2/0.8-0.9. 0.2-0.3/0.7-0.8, 0.3-0.4/0.6-0.7, 0.4-0.5/0.5-0.6). A 5 bias x 3 drug RM-ANOVA was used to compare mean baseline pupil diameter across these five belief levels and between drug-groups.

6.3.11.3 Pupil diameter modulation by predictions and beliefs

For an initial model-agnostic assessment of pupil diameter modulation across post-response and post-outcome periods, trials were binned according to whether participants had made a correct or incorrect prediction about outcome type. In my previous work with de Berker et al., it was possible to show that pupil responsivity to probabilistic outcomes is influenced by participants' beliefs about surprise and irreducible uncertainty (de Berker et al., 2016). Therefore, I next implemented median splits to separate trials according to whether they were high or low in participant-specific surprise (model parameter $|\delta_1|$) and irreducible uncertainty ($\hat{\mu}_1$).

Note that $\hat{\mu}_1$ reflects participants' estimated beliefs of the current cue:outcome probabilities. These estimates lie between 0 and 1 and encapsulate two bias directions, e.g., a $\hat{\mu}_1$ estimate of 0.1 indicates an equivalent bias magnitude as a $\hat{\mu}_1$ estimate of 0.9, but in the opposite bias direction. Maximal irreducible uncertainty occurs at p=0.5. Since $\hat{\mu}_1$ showed a symmetrical inverted-U relationship with baseline pupil diameter (see section 6.4.3.1), the median splits were conducted using a $\hat{\mu}_1$ parameter that had been adjusted to scale between 0 and 0.5. As such, any raw $\hat{\mu}_1$ values between 0 and 0.5 remained unchanged, while those between 0.5 and 1 were transformed as follows:

$$Adjusted\ \hat{\mu}_1 = -\hat{\mu}_1 + 1$$

Equation 6.1

In accordance with the prediction that NA would influence phasic volatility estimates (μ_3) and/or phasic volatility uncertainty ($\hat{\sigma}_3$), I also performed additional median splits on high and low μ_3 and $\hat{\sigma}_3$. Note that $\hat{\sigma}_3$ captures an individual's uncertainty about their current phasic volatility estimate, in contrast to volatility uncertainty which is the uncertainty that arises due to environmental instability.

RM-ANOVAs were applied to assess peak pupil diameter during the post-response and post-outcome periods for these different trial-types and across drug-groups. Peak pupil

responses were calculated according to the method described in section 6.3.11.1. Where significant drug effects were identified, additional exploratory analyses were conducted on the individual drug-groups to further characterise the effect of beliefs on pupil diameter. A Benjamini-Hochberg correction was applied to each of these additional analyses to control for FDR. No analyses were conducted on the post-cue period as it was fixed at 300ms to minimise overlap with the timing of the button-press responses (Table 6.4).

6.3.11.4 Regression analyses

To extend these analyses by pinpointing the specific effects of participants' beliefs on pupil diameter, and assessing the impact of the noradrenergic manipulations on these pupil responses, a regression approach implemented by Browning et al. was adopted (Browning et al., 2015). Specifically, regression analyses were conducted to examine the effects of trial-wise estimates of surprise, irreducible uncertainty, phasic volatility, and phasic volatility uncertainty on pupil dilation during the post-response and post-outcome periods. For the Placebo group, pupil diameter during the post-response period (0 to 1s from response onset) was sampled using 120 8.3ms bins (i.e., 1s/sample rate). Pupil diameter during the post-outcome period (0 to 3.3s from outcome onset) was sampled using 396 8.3ms bins. The total duration of the post-outcome period entered into the regression analyses was the minimum that could occur on any given trial.

Regression analyses were conducted for each of these bins, with surprise ($|\delta_1|$), irreducible uncertainty ($\hat{\mu}_1$), phasic volatility estimate (μ_3), and phasic volatility uncertainty ($\hat{\sigma}_3$) entered as regressors of interest. As in section 6.3.11.3, the adjusted irreducible uncertainty measure, which peaks at p=0.5 (see Equation 6.1), was used in all regression analyses.

Cue type (-1 for low-pitch tone, 1 for high-pitch tone), response type (-1 for left button-press, 1 for right button-press), and outcome type (-1 for "cow", 1 for "pig") were entered into the analysis as control regressors.

For each participant in each of the active drug-groups, the mean Placebo regression beta weights across the post-response and post-outcome periods were subtracted from the time-series of beta weights for that participant.

The resulting time-series of beta weights (for the Placebo group) and Δbeta-weights (for the active drug-groups) for the constant component of pupil diameter, surprise, irreducible uncertainty, phasic volatility estimate, and phasic volatility uncertainty were down-sampled to give mean beta weight estimates of the effects of each factor on pupil

dilation for 6 sequential 166.7ms bins across the post-response period and 19 sequential 167.7ms bins across the post-outcome period.

For the Placebo group, t-tests were used to determine whether each down-sampled beta-weight bin differed significantly from zero. Each active drug-group was compared to Placebo by using t-tests to determine whether each down-sampled Δbeta-weight bin differed significantly from zero. Since 6 comparisons were made across the post-response period, 19 comparisons were made across the post-outcome period, and these comparisons were made across three drug-groups, a Benjamini-Hochberg correction was applied to correct for the FDR arising from the total of 75 comparisons.

6.3.12 Behaviour vs pupil responses

To assess whether the pupillary responses to participants' beliefs were associated with altered behaviour during the PLT, correlation analyses were conducted. Specifically, for each of the three drug-groups, a Pearson's correlation was used to compare the mean volatility beta weight from 0-1.7s post-outcome, and the mean irreducible uncertainty beta weight from 0.7-1.7s post-outcome, to participants' mean learning rate and mean performance score.

6.3.13 Control analyses

6.3.13.1 Model parameter correlations

To verify that the output of the regression analyses was not complicated by high correlations between the model parameters entered as regressors, correlations between $|\delta_1|$, $\hat{\mu}_1$, μ_3 and $\hat{\sigma}_3$ were assessed.

6.4 Results

Behavioural data for 90 participants are reported. The three groups were matched for gender (Kruskal-Wallis test: H_2=0.00, p=1.000), age (one-way ANOVA: $F_{2,89}$=1.29, p=0.281), body weight ($F_{2,89}$=0.082, p=0.921), education level (H_2=4.79, p=0.091), and all other baseline psychometric measures taken (Table 6.2). Pupil data for one participant from the NA+ group is missing due to a technical problem with the eyetracker at the time of recording.

6.4.1 Model-agnostic results

On average, 45.9 ± 1.28% (±SEM), 47.9 ±1.28% and 46.5 ± 1.28% of predictions made during the unbiased (0.5/0.5) blocks were correct in the NA-, Placebo and NA+ groups

respectively, indicating that participants were performing close to chance level when there was no cue:outcome bias (Figure 6.3A). A one-way ANOVA indicated that accuracy (i.e., the percentage of correct responses) in the unbiased blocks did not differ between drug-groups ($F_{2,86}$=0.55, p=0.581). Similarly, decision times during unbiased blocks were equivalent across drug-groups ($F_{2,86}$=0.36, p=0.697; Figure 6.4A), indicating that the drug manipulations did not merely modulate participants' ability to respond quickly. Neither effect was modulated by Δalertness (both p≥0.74).

	Placebo (n = 30)	**NA- (n = 30)**	**NA+ (n = 30)**	**Between-groups difference?**
Gender *(number male)*#	13	13	13	ns p = 1.000
Age *(years)*	24.1 ± 4.2	26.0 ± 5.9	24.8 ± 3.9	ns p = 0.281
Weight *(kg)*	69.5 ± 2.8	68.3 ± 2.2	68.5 ± 1.8	ns p = 0.921
Education Level *(1-5)*#	2.8 ± 0.2	3.1 ± 0.1	3.4 ± 0.2	ns p = 0.091
Digit Span *(forwards + backwards)*#	12.8 ± 0.5	12.8 ± 0.4	13.4 ± 0.4	ns p = 0.334
Impulsivity: *BIS-11*	64.4 ± 2.0	61.5 ± 1.4	61.4 ± 1.4	ns p = 0.337
Risk-taking: *DOSPERT (total)*	113.1 ± 3.6	104.3 ± 4.0	108.2 ± 3.8	ns p = 0.699
Distractibility: *CFQ*	39.5 ± 2.0	41.4 ± 2.6	40.0 ± 2.2	ns p = 0.833
Sleep quantity on the previous night *(hours)*#	7.0 ± 0.2	7.6 ± 0.3	7.1 ± 0.2	ns p = 0.107
Sleep quality on the previous night *(1-8)*#	5.2 ± 0.2	5.4 ± 0.2	5.5 ± 0.3	ns p = 0.615
Fatigue during task *(0 – 100)*	45.4 ± 3.9	44.7 ± 3.8	46.6 ± 4.1	ns p = 0.941
Active drug *(%)*#	37	57	73	p = 0.017

Table 6.2 Summary details for participants in each experimental group. *Between-groups comparisons revealed no significant differences (ns = non-significant) for gender, age, body weight, education level, baseline working memory (Digit Span), impulsivity (Barratt Impulsiveness Scale; BIS-11), risk-taking (Domain-Specific Risk-Taking Scale; DOSPERT), distractibility (Cognitive Failures Questionnaire; CFQ), fatigue during the task, or sleep quality or quantity on the previous night. For continuous data, one-way ANOVAs were used to test for any between-group differences. For discrete data (#), Kruskal-Wallis tests were applied. Education Level refers to the highest attained from the following: 1 = compulsory education (≤ 12 years); 2 = further education (13-14 years); 3 = undergraduate degree (15-17 years); 4 = one postgraduate degree (≥ 18 years); 5 =*

multiple postgraduate degrees. Age data are mean ± SD. Remaining data are mean ± SEM. Active drug refers to the percentage of participants within each group who reported at the end of the experiment that they believed they had received an active drug.

6.4.1.1 Accuracy increases with increasing cue:outcome bias

A 3 bias (high/moderate/none) x 2 time (Early/Late trials) x 3 drug RM-ANOVA revealed that accuracy increased significantly as the true cue:outcome bias increased (effect of bias: $F_{1.70,146.55}$=337.45, p<0.001, η_p^2=0.80; Figure 6.3B). Additional (FDR-corrected) exploratory analyses revealed that the effect of bias existed in all three drug-groups (NA-: $F_{1.58,44.17}$=155.68, p<0.001, η_p^2=0.85; Placebo: $F_{1.66,46.44}$=118.37, p<0.001, η_p^2=0.81; NA+: $F_{2,56}$=81.65, p<0.001, η_p^2=0.75). Together with the significant increase in accuracy over the course of a contextual block (effect of time $F_{1,86}$=20.58, p<0.001, η_p^2=0.19), this indicates that participants learned to estimate the true cue:outcome probabilities. The 3 bias x 2 time x 3 drug RM-ANOVA indicated a significant bias x time interaction ($F_{1.83,157.39}$=7.80, p=0.001, η_p^2=0.08), but none of the main effects were modulated by drug (all p>0.21) or Δalertness (all p>0.13).

Since learning would not be expected over the course of an unbiased block, I examined whether the bias x time interaction was driven by an increase in accuracy during the biased, but not the unbiased, blocks. A 2 time x 3 drug RM-ANOVA on the percentage correct responses in the unbiased blocks indeed revealed no effect of time (p=0.604) and no modulation by drug (p=0.808) or Δalertness (p=0.940). In contrast, a 2 bias x 2 time x 3 drug RM-ANOVA on the highly and moderately biased blocks revealed significant effects of bias ($F_{1,86}$=437.97, p<0.001, η_p^2=0.84) and time ($F_{1,86}$=41.20, p<0.001, η_p^2=0.32), and no modulation by drug (all p>22) or Δalertness (all p>0.09). The increase in accuracy over time was equivalent across highly and moderately biased blocks (no probability x time interaction: p=0.168).

Further (FDR-corrected) exploratory analyses indicated that accuracy increased across the course of the highly biased blocks in all three drug-groups (NA-: $F_{1,28}$=9.94, p=0.004, η_p^2=0.26; Placebo: $F_{1,28}$=36.87, p<0.001, η_p^2=0.57; NA+: $F_{1,28}$=5.76, p=0.023, η_p^2=0.17), and across the course of the moderately biased blocks in the NA- ($F_{1,28}$=5.85, p=0.022, η_p^2=0.17) and Placebo groups ($F_{1,28}$=10.281, p=0.003, η_p^2=0.27). NA+ group accuracy remained unchanged across the course of the moderately biased blocks (p=0.627).

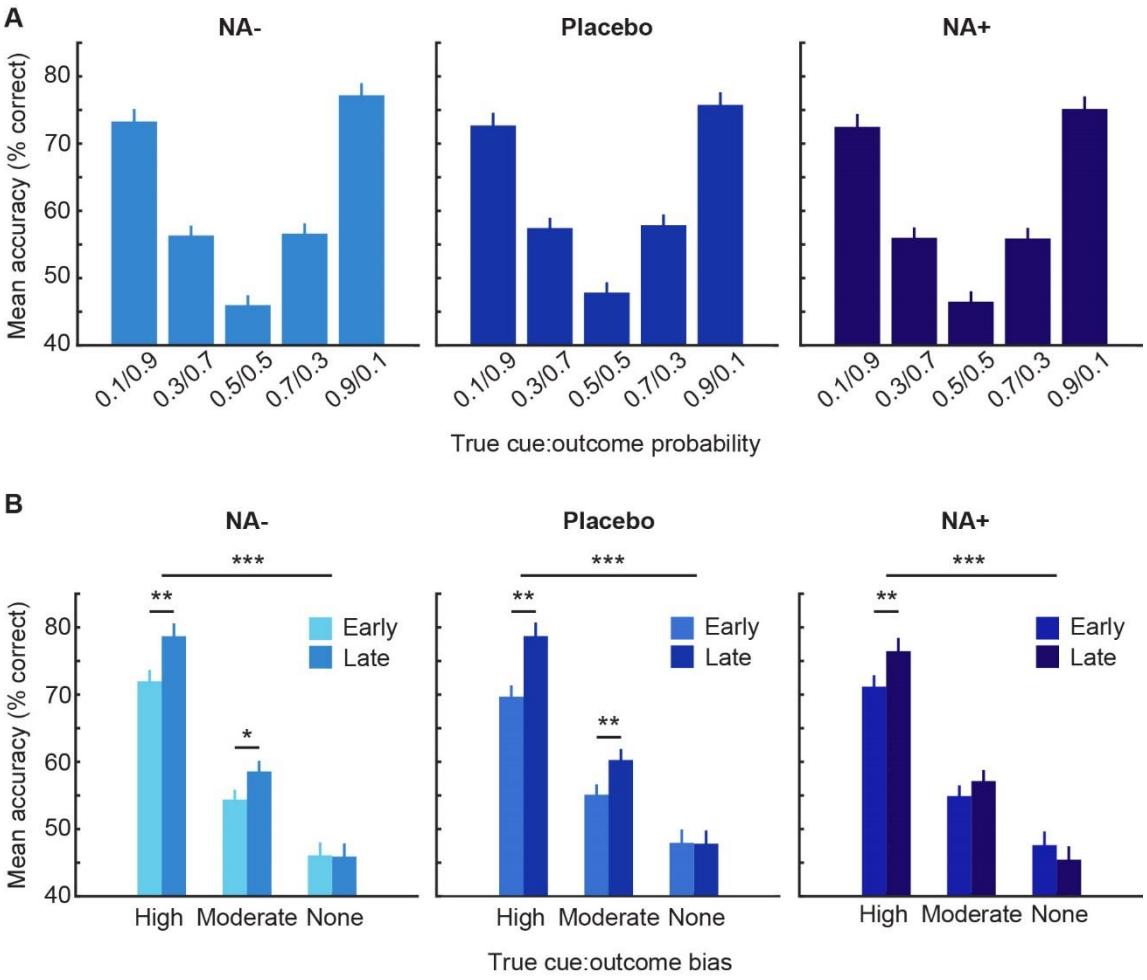

Figure 6.3 Model-agnostic analysis of accuracy. *Participants in all three drug-groups demonstrated learning of the underlying the cue:outcome relationships. (A) Across groups, participants made correct predictions on a higher percentage of trials as the true cue:outcome bias increased. (B) This learning was observed over the course of the biased contextual blocks, with participants in all three drug-groups achieving a higher accuracy on trials in the second half of the highly biased blocks. Accuracy rates in the NA- and Placebo groups also increased over the course of the moderately biased blocks. Accuracy remained unchanged over the course of the unbiased blocks in all three drug-groups. Results are mean ± SEM, corrected for Δalertness. * p<0.05, ** p<0.01, *** p<0.001, after an FDR correction for three multiple comparisons.*

6.4.1.2 Decision times decrease with increasing cue:outcome bias

A 3 bias (high/moderate/none) x 2 time (Early/Late trials) x 3 drug RM-ANOVA indicated that participants' decision times decreased significantly as the true cue:outcome bias increased (effect of bias: $F_{2,172}$=17.96, p<0.001, η_p^2=0.17; Figure 6.4B), again indicative of learning of the true cue:outcome probabilities. Additional (FDR-corrected) exploratory analyses revealed that the effect of bias existed in the NA- ($F_{2,56}$=11.50, p<0.001,

η_p^2=0.29) and Placebo ($F_{2,56}$=5.58, p=0.006, η_p^2=0.17) groups, and there was a trend-level effect in the NA+ group ($F_{2,56}$=3.15, p=0.051, η_p^2=0.10). However, there was no significant change in decision time over the course of the contextual blocks (no effect of time: p=0.164). The 3 bias x 2 time x 3 drug RM-ANOVA indicated that the effect of bias was not modulated by drug (p=0.173) or Δalertness (p=0.708). However, there was a significant time x drug interaction ($F_{2,86}$=4.24, p=0.018, η_p^2=0.09).

Again, since there was a bias x time interaction ($F_{2,72}$=6.64, p=0.002, η_p^2=0.07), and because a decrease in decision time would not necessarily be expected across the course of the unbiased blocks, a 2 time x 3 drug RM-ANOVA was conducted on the decision times in the unbiased blocks. This indeed revealed no significant effect of time (p=0.155), and no modulation by drug (p=0.127) or Δalertness (p=0.183). In contrast, a 2 bias x 2 time x 3 drug RM-ANOVA on the highly and moderately biased blocks revealed significant effects of bias ($F_{1,86}$=27.63, p<0.001, η_p^2=0.24) and time ($F_{1,86}$=16.09, p<0.001, η_p^2=0.16), and a significant time x drug interaction ($F_{2,86}$=3.18, p=0.047, η_p^2=0.07). Post-hoc (FDR-corrected) pairwise comparisons between the active drug-groups and Placebo indicated that the time x drug interaction was driven by the NA+ group, with participant-specific differences in decision times on Late vs Early trials demonstrating less speeding across biased blocks compared to Placebo (t_{56}=2.52, p=0.014, Cohen's *d*=0.67). There was a trend-level bias x drug interaction ($F_{2,86}$=3.05, p=0.053, η_p^2=0.07), but none of the effects were modulated by Δalertness (all p>0.41).

Further (FDR-corrected) exploratory analyses indicated that decision time decreased across the course of the highly biased blocks in the NA- ($F_{1,28}$=12.139, p=0.002, η_p^2=0.30) and Placebo ($F_{1,28}$=16.304, p<0.001, η_p^2=0.37) groups, but not in the NA+ group (p=0.693). Uncorrected analyses indicated that Placebo group decision times also decreased over the course of the moderately biased blocks ($F_{1,28}$=5.93, p=0.022, η_p^2=0.18), but the result did not survive an FDR correction for multiple comparisons. No change in decision times across the course of the moderately biased blocks was observed in the NA- or NA+ groups (both p>0.88).

In summary, the increasing accuracy and decreasing decision time that accompanied an increase in cue:outcome bias, and continued over the course of the biased blocks, indicates that participants in all three drug-groups demonstrated learning of the task's underlying probabilistic structure. Moreover, the smaller decrease in decision times across biased blocks in the NA+ group compared to Placebo is indicative of a modulation of learning by reboxetine.

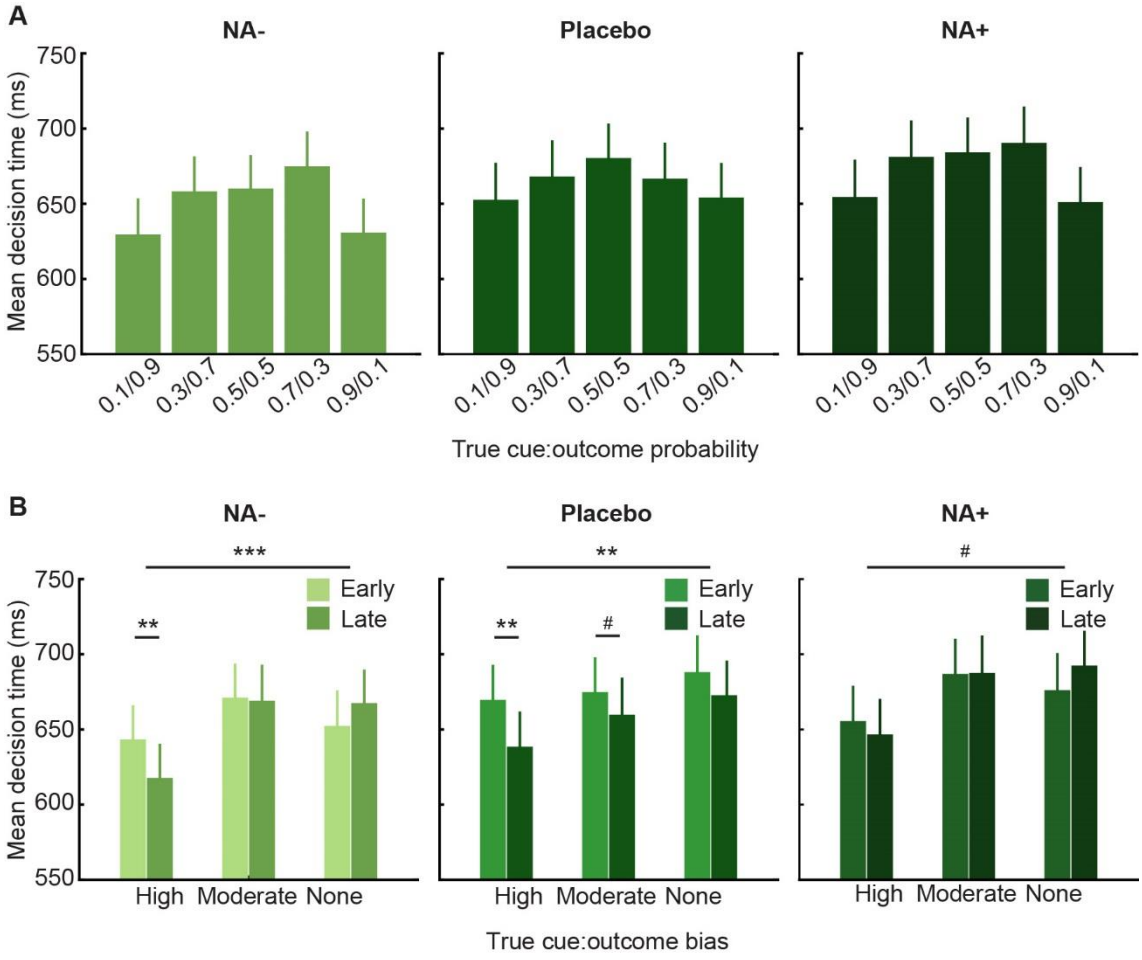

Figure 6.4 Model-agnostic analysis of decision times. *(A) Decision times decreased as the true cue:outcome bias increased, further demonstrating learning of the task's structure in the three drug-groups. (B) Learning was observed over the course of the biased contextual blocks, with participants in the NA- and Placebo groups showing decreased decision times on trials in the second half of the highly biased blocks. A trend-level decrease in decision time was also observed across the course of the moderately biased blocks in the Placebo group. Decision times remained unchanged over the course of the unbiased blocks in all three drug-groups. Results are mean ± SEM, corrected for Δalertness. * p<0.05, ** p<0.01, *** p<0.001, # trend, after an FDR correction for three multiple comparisons.*

6.4.1.3 Task performance is modulated by cue:outcome bias and noradrenaline

Next, accuracy and decision time were combined to calculate a performance score (Figure 6.5). A 3 bias x 2 time x 3 drug RM-ANOVA on performance scores demonstrated that performance improved significantly as the true cue:outcome bias increased (effect of probability: $F_{1.66,142.92}=252.41$, $p<0.001$, $\eta_p^2=0.75$) and as a contextual block progressed (effect of time: $F_{2,86}=23.73$, $p<0.001$, $\eta_p^2=0.22$). There was a significant bias

x time interaction ($F_{2,172}$=15.50, p<0.001, η_p^2=0.153) and a trend-level time x drug interaction ($F_{2,86}$=3.08, p=0.051, η_p^2=0.07). Post-hoc pairwise (FDR-corrected) comparisons on the participant-specific differences in performance scores on Late vs Early trials demonstrated that this interaction was driven by the NA+ group (Figure 6.5C). Indeed, performance was poorer in the NA+ group compared to Placebo (t_{56}=2.67, p=0.016, *d*=0.71). In the 3 bias x 2 time x 3 drug RM-ANOVA, neither the effect of bias nor the bias x time interaction was modulated by drug (both p>0.52). None of the effects were modulated by Δalertness (all p>0.09).

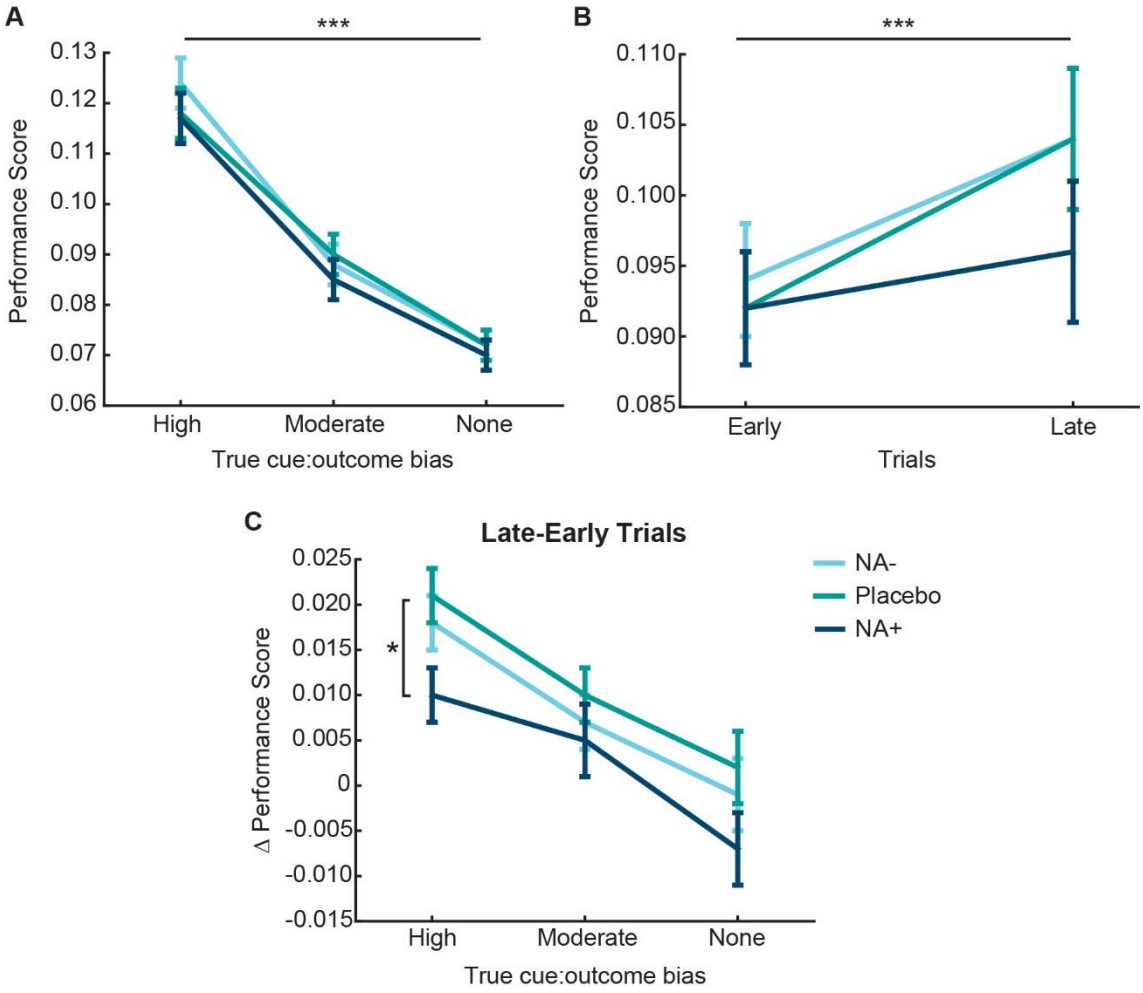

Figure 6.5 Model-agnostic analysis of performance scores. *In each case, a higher score indicates better performance since it reflects a higher accuracy and/or faster decision time. (A) Performance improved significantly in each drug-group as the true cue:outcome probability increased. (B) Performance also improved across the course of the contextual blocks with subjects achieving higher performance scores on Late trials (i.e., those in the second half of the contextual blocks) compared to Early trials (i.e., those in the first half of the contextual blocks). Here the data have been collapsed across the three bias levels. (C) Assessment of the participant-specific differences in performance*

*scores on Late vs Early trials across the three bias levels indicates that performance was poorer in the NA+ group compared to Placebo. Data are mean ± SEM, corrected for Δalertness. * p<0.05, ** p<0.01, *** p<0.001.*

6.4.2 Model-based results

6.4.2.1 Decision times are modulated by irreducible uncertainty

Participants' estimated beliefs about the current cue:outcome probabilities, and hence their current irreducible uncertainty (model parameter $\hat{\mu}_1$), predicted their decision times (Figure 6.6A). Indeed, a curve describing the variance of a Bernoulli distribution representing beliefs about probabilities, which is the formulation of irreducible uncertainty $\hat{\mu}_1$ used here, predicts mean decision times across the three drug-groups (Pearson's correlations for NA: r=0.942, p<0.001; Placebo: r=0.902, p<0.001; NA+: r=0.931, p<0.001), replicating my previous work with de Berker et al. (de Berker et al., 2016)

A 5 bias x 3 drug RM-ANOVA indicated that participants' decision times increased significantly as their estimates of cue:outcome bias decreased, and thus as their estimated irreducible uncertainty increased (effect of bias: $F_{2.38,188.25}$=60.69, p<0.001, η_p^2=0.43; Figure 6.6B). Additional (FDR-corrected) exploratory analyses revealed that the effect of bias existed in all three drug-groups (NA-: $F_{1.88,48.75}$=19.48, p<0.001, η_p^2=0.43; Placebo: $F_{2.50,64.98}$=36.69, p<0.001, η_p^2=0.59; NA+: $F_{2.89,72.22}$=13.97, p<0.001, η_p^2=0.36). Indeed, the effect of estimated bias on decision times was not modulated by drug (p=0.723), or by Δalertness (p=0.310).

Since the HGF is informed of participants' trial-wise predictions, but not their decision times, this increase in decision time with decreasing bias and increasing irreducible uncertainty (i.e., as beliefs approach p=0.5) indicates that the model captured participants' beliefs well. Indeed, the relationship between decision time and probability is clearer when trials are categorised according to participants' estimated beliefs about the current cue:outcome probability compared to when trials are categorised according to the true cue:outcome probabilities, which were hidden from participants (c.f. Figure 6.4).

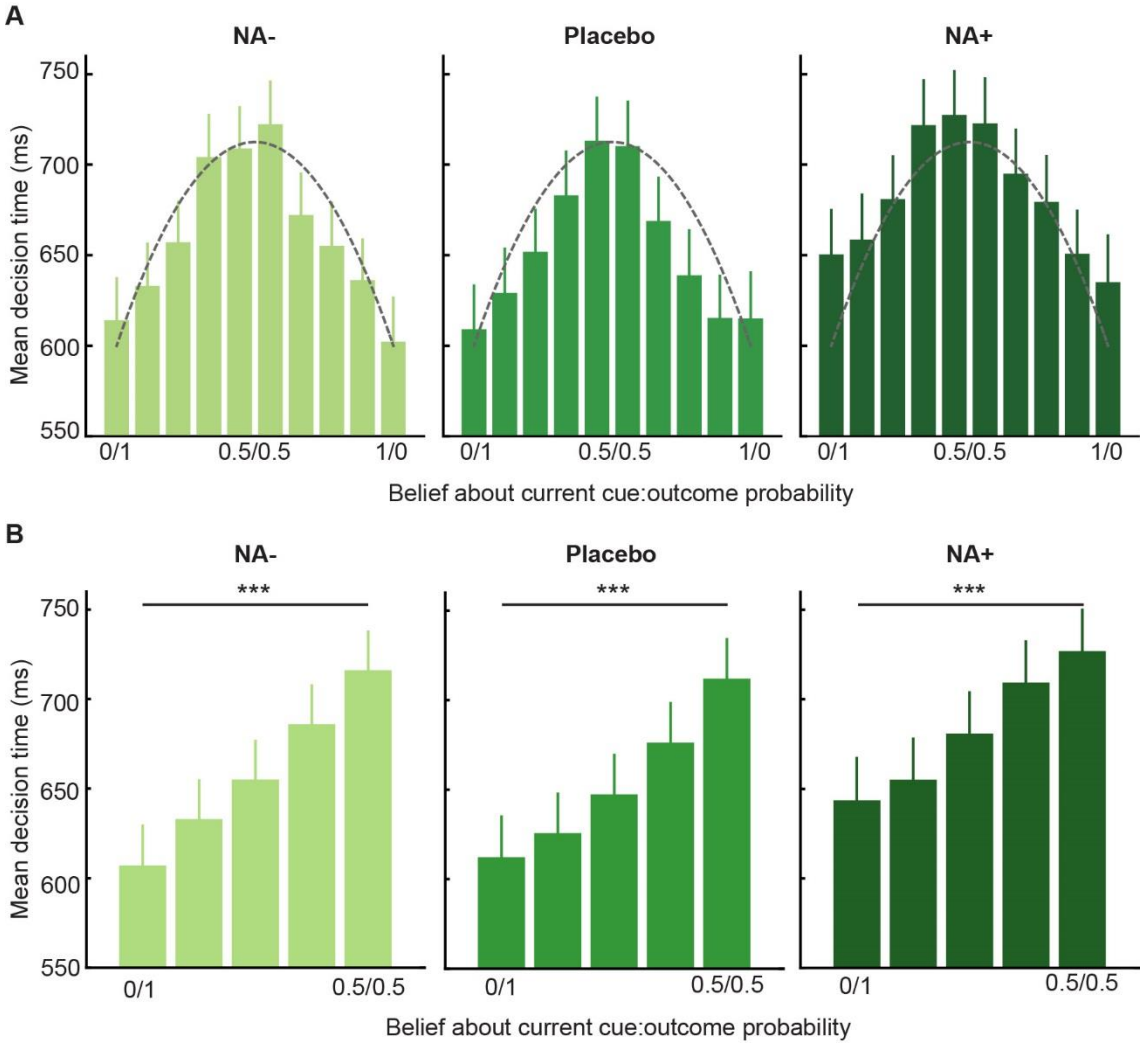

Figure 6.6 Decision times according to participants' cue:outcome probability beliefs. *(A) Participants' estimated beliefs about cue:outcome probabilities (and their beliefs about irreducible uncertainty) lay between 0 and 1. Across drug-groups, the distribution of decision times across these estimates conformed closely to a Bernoulli distribution (grey dashed line), with decision times peaking with maximal irreducible uncertainty ($\hat{\mu}_1$=0.5). (B) For statistical analysis of the effect of irreducible uncertainty on decision times, bins showing an equivalent bias magnitude (but in opposite directions) were collapsed, to create 5 bias bins spanning 0/1 to 0.5/0.5. Across the three drug-groups, decision time increased significantly as irreducible uncertainty about the cue:outcome relationships increased. Since the HGF is uninformed of participants' decision times, this relationship with estimated irreducible uncertainty indicates that the model captured participants' beliefs well. Data are mean ± SEM, corrected for Δalertness. *** p<0.001, after an FDR correction for three comparisons.*

6.4.3 Pupil analyses

6.4.3.1 Baseline pupil diameter is modulated by irreducible uncertainty

Baseline pupil diameter on each trial displayed a clear inverted-U relationship with belief about the current cue:outcome probabilities, and thus with estimates of irreducible uncertainty, as reflected by model parameter $\hat{\mu}_1$ (Figure 6.7A). This recapitulates the relationship between decision times and irreducible uncertainty, and replicates my previous work with de Berker et al. (de Berker et al., 2016). A curve describing the variance of a Bernoulli distribution representing beliefs about cue:outcome probabilities predicts mean baseline pupil diameter extremely well across the three drug-groups (Pearson's correlations for NA: r=0.886, p<0.001; Placebo: r=0.921, p<0.001; NA+: r=0.938, p<0.001).

A 5 bias x 3 drug RM-ANOVA indicated that participants' baseline pupil diameter increased significantly as their estimates of cue:outcome bias decreased and therefore as their estimated irreducible uncertainty increased (effect of bias: $F_{1.91,149.25}$=16.98, p<0.001, η_p^2=0.18; Figure 6.7B). Additional (FDR-corrected) exploratory analyses revealed that the effect of bias existed in all three drug-groups (NA-: $F_{1.98,51.49}$=5.68, p<0.001, η_p^2=0.18; Placebo: $F_{1.83,47.49}$=6.17, p<0.001, η_p^2=0.19; NA+: $F_{1.86,44.59}$=9.83, p<0.001, η_p^2=0.29). Indeed, the effect of estimated bias on baseline pupil diameter was not modulated by drug (p=0.613), or by Δalertness (p=0.564).

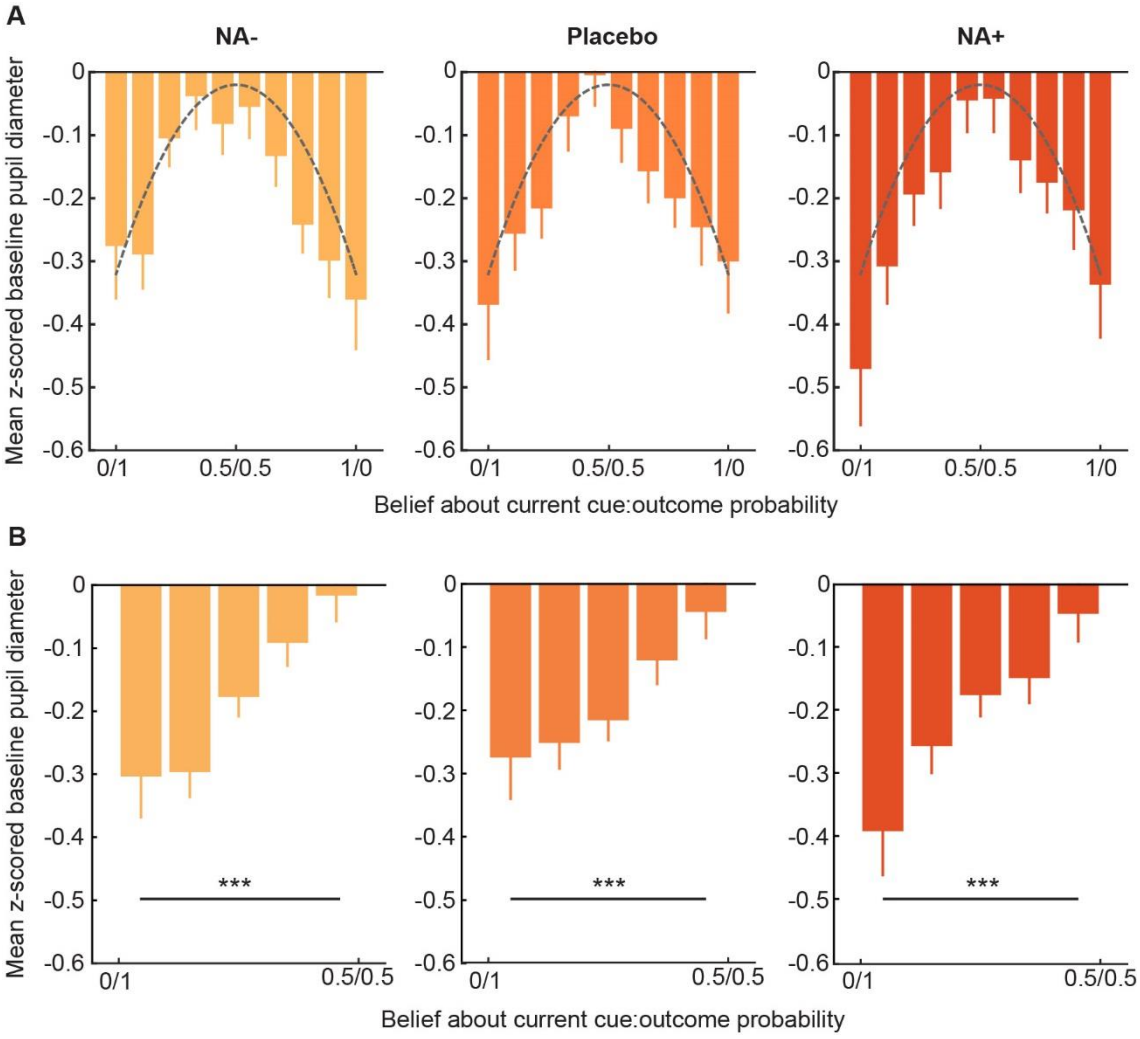

Figure 6.7 Baseline pupil diameter according to participants' cue:outcome probability beliefs. *(A) Across drug-groups, the relationship between pupil diameter and beliefs about cue:outcome probabilities (and thus irreducible uncertainty) closely conformed to a Bernoulli distribution (grey dashed line), with pupil diameters peaking with maximal irreducible uncertainty ($\hat{\mu}_1$=0.5). (B) For statistical analysis of the effect of irreducible uncertainty on pupil diameter, bins showing an equivalent bias magnitude (but in opposite directions) were collapsed to create 5 bias bins spanning 0/1 to 0.5/0.5. Across the three drug-groups, pupil diameter increased significantly as irreducible uncertainty about the cue:outcome relationships increased. Data are mean ± SEM, corrected for Δalertness. *** p<0.001, after an FDR correction for three comparisons.*

6.4.3.2 Event-related analysis of pupil diameter

By epoching pupil diameter during the PLT by trial, and then according to cue, response and outcome onset within each trial, it was possible to characterise pupil diameter modulations during the post-response and post-outcome periods (Figure 6.8). Across

drug-groups, pupil diameter started to increase after cue onset, and continued to increase following the button-press response participants made to indicate their prediction about which outcome would follow. Following outcome onset, pupils showed a positive evoked response. The peak post-response increase in pupil diameter was identical across drug-groups ($p=0.802$). Compared to Placebo, the peak post-outcome increase in pupil diameter tended to be augmented in the NA+ group and suppressed in the NA- group. However, a one-way ANOVA on the post-outcome pupillary peak revealed no significant modulation by drug-group ($p=0.202$). No reported effects were modulated by Δalertness ($p \geq 0.37$).

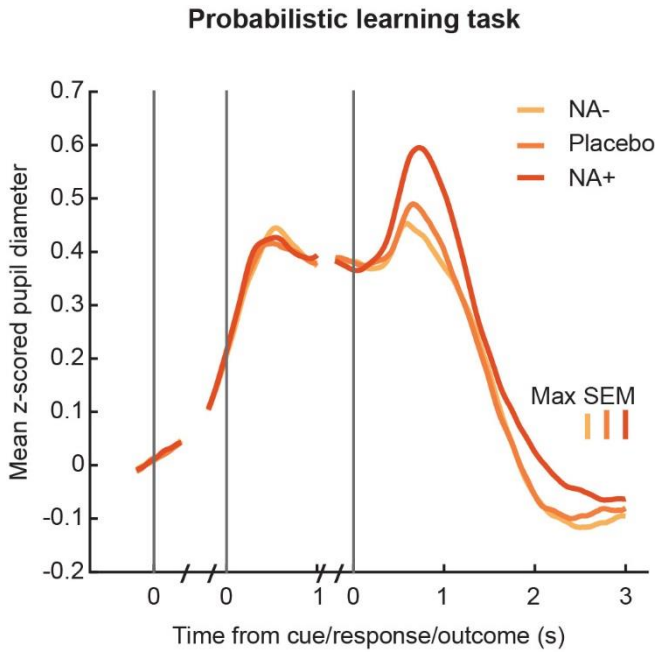

Figure 6.8 Trial-wise pupil diameter during the probabilistic learning task (PLT). *Data have been baseline-corrected and epoched according to (1) cue, (2) response and (3) outcome onset. Across the three drug-groups, pupil diameter begins to increase after cue onset, continues to increase following the button-press response participants make to indicate their prediction about which outcome will follow, and then shows a positive evoked response following outcome onset. Compared to Placebo, there was a tendency for the post-outcome increase in pupil diameter to be augmented in the NA+ group and suppressed in the NA- group. Data are mean pupil diameter across participants. Error bars indicate the maximum SEM for each drug-group.*

Applying the same epoching approach to the control data, pupil diameter was again found to start increasing after cue onset, and to continue increasing following the participants' button-press response (Figure 6.9). Unlike during the PLT, there was no positive evoked response following outcome presentation in any of the drug-groups.

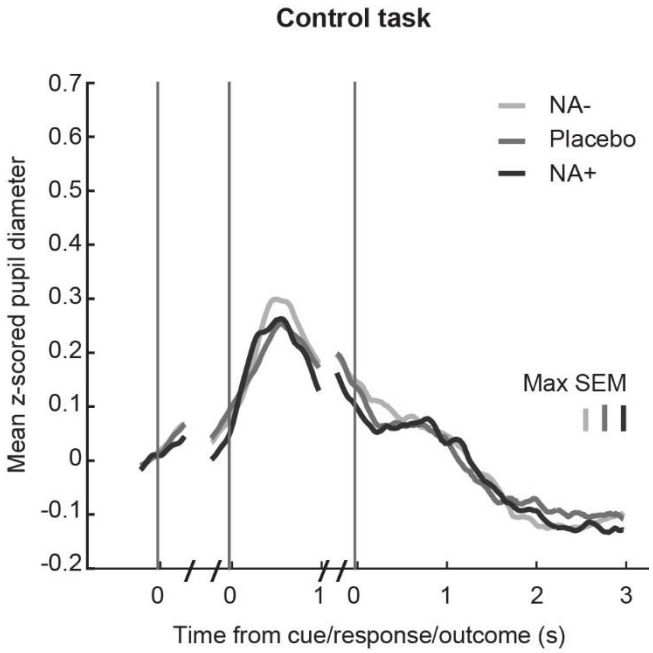

Figure 6.9 Trial-wise pupil diameter during the control task (CT). *Data have been baseline-corrected and epoched according to (1) cue, (2) response and (3) outcome onset. Across the three drug-groups, pupil diameter begins to increase after cue onset, and continues to increase following the button-press response participants make to indicate their "prediction" about which outcome will follow. Unlike during the PLT, there is no positive evoked response following outcome presentation. Data are mean pupil diameter across participants. Error bars indicate the maximum SEM for each drug-group.*

The key difference between the PLT and the CT was that the latter was free from uncertainty about which outcome would follow the cue on any given trial. Therefore, the fact that no increase in pupil diameter was observed post-outcome in the CT suggests that the evoked response observed during the PLT was indeed due to uncertainty and/or surprise about trial-wise outcome presentation.

In both the PLT and CT, an increase in pupil diameter was observed following the button-press response participants made to indicate their prediction about which outcome would follow. Given that this pupillary response was observed under both conditions of uncertainty and conditions of no uncertainty about outcome type, this suggests that the increase in pupil diameter at this stage in the trial is linked to the act of making a decision by button-press. Moreover, across drug-groups, the maximum pupil diameter during the post-response period was greater during the PLT (mean for NA-: 0.445; Placebo: 0.416; NA+: 0.427) than the CT (mean for NA-: 0.299; Placebo: 0.253; NA+: 0.263), and was sustained for longer, suggesting that uncertainty during the PLT may have had an

additional dilatory effect on pupils over and above the effect of making a button-press response.

6.4.3.3 Model-agnostic analyses of pupil diameter modulation across post-response and post-outcome periods

A RM-ANOVA indicated that the peak increase in pupil diameter following outcome presentation was greater on trials on which participants had made an incorrect prediction about outcome type compared to when they had made a correct prediction ($F_{1,85}$=29.91, p<0.001, η_p^2=0.26; Figure 6.10). However, this effect was modulated by drug-group ($F_{2,85}$=29.91, p=0.002, η_p^2=0.14). Indeed, additional exploratory analyses demonstrated that this increased pupillary response on incorrect trials existed in the NA- ($F_{1,28}$=11.17, p=0.003, η_p^2=0.29) and NA+ ($F_{1,27}$=18.22, p<0.001, η_p^2=0.40) groups, but not in the Placebo group (p=0.348). None of these effects were modulated by Δalertness (p>0.34). In contrast, there was no effect of drug on post-response pupillary dilation (p=0.224).

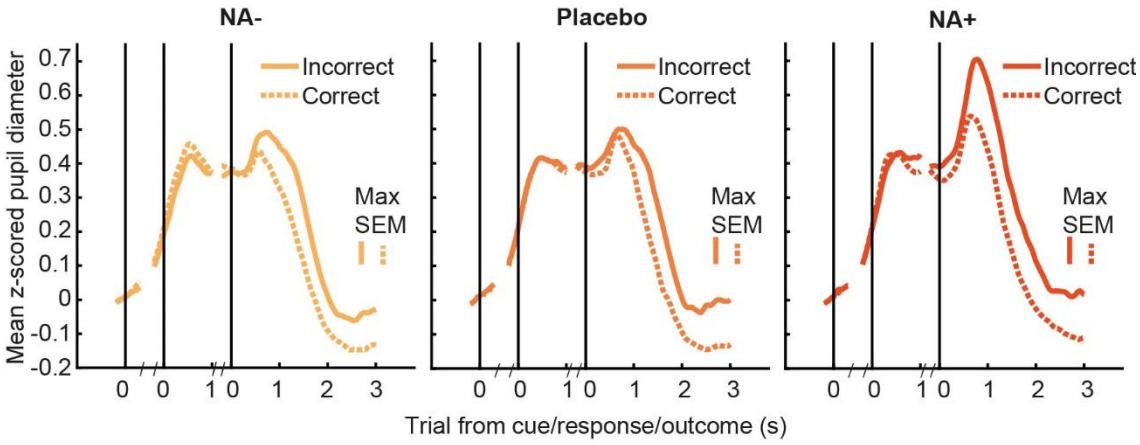

Figure 6.10 Pupil diameter on trials with correctly and incorrectly predicted outcomes. *An incorrect prediction augmented the post-outcome dilatory pupillary response in the NA- and NA+ groups. Data are mean pupil diameter across participants. Error bars indicate the maximum SEM for each trial-type.*

6.4.3.4 Model-based analyses of pupil diameter modulation across post-response and post-outcome periods

Using estimates of participants' trial-wise beliefs, as provided by the HGF, to categorise trials indicated that high surprise ($|\delta_1|$) and high irreducible uncertainty ($\hat{\mu}_1$) were associated with an increase in the pupillary response observed in the post-outcome period compared to low surprise ($F_{1,85}$=45.52, p<0.001, η_p^2=0.35; Figure 6.11A) and low irreducible uncertainty ($F_{1,85}$=45.15, p<0.001, η_p^2=0.35; Figure 6.11B), respectively. The positive effect of surprise on post-outcome pupil diameter is reassuring given the positive

effect of incorrect responses observed previously (Figure 6.10), and given that the surprise quantity estimated by the HGF is correlated with correct/incorrect predictions (NA-: mean r=0.589 ± 0.03 (± SEM); Placebo: r=0.611 ± 0.03; NA+: r=0.562 ± 0.04).

These effects were equivalent across drug-groups (all p>0.50). There was also a significant effect of high irreducible uncertainty during the post-response period ($F_{1,85}$=7.05, p=0.009, η_p^2=0.08), but this effect was not significant within the individual NA- (p=0.465), Placebo (p=0.465) or NA+ (p=0.054) groups.

A RM-ANOVA with drug-group as a between-subjects factor indicated that high phasic volatility estimates (μ_3) were associated with an increased post-outcome pupillary response ($F_{1,85}$=11.89, p=0.001, η_p^2=0.12; Figure 6.11C). Additional exploratory analyses indicated that this effect was driven by the NA- group ($F_{1,28}$=11.49, p=0.006, η_p^2=0.29), but not the Placebo (p=0.110) or NA+ (p=0.433) groups. Similarly, high phasic volatility uncertainty estimates ($\hat{\sigma}_3$) were associated with an increased post-outcome pupillary response when drug-group was included in a RM-ANOVA as a between-subjects factor ($F_{1,85}$=6.81, p=0.011, η_p^2=0.07; Figure 6.11D), but additional analyses on the individual drug-groups demonstrated that this effect was not significant within the NA- (p=0.033), Placebo (p=0.259) or NA+ (p=0.141) groups. There was also a significant effect of high phasic volatility uncertainty during the post-response period ($F_{1,85}$=5.98, p=0.021, η_p^2=0.18), but again this effect was not significant within the individual drug-groups (all p≥0.058). None of the reported effects were modulated by Δalertness (all p>0.09).

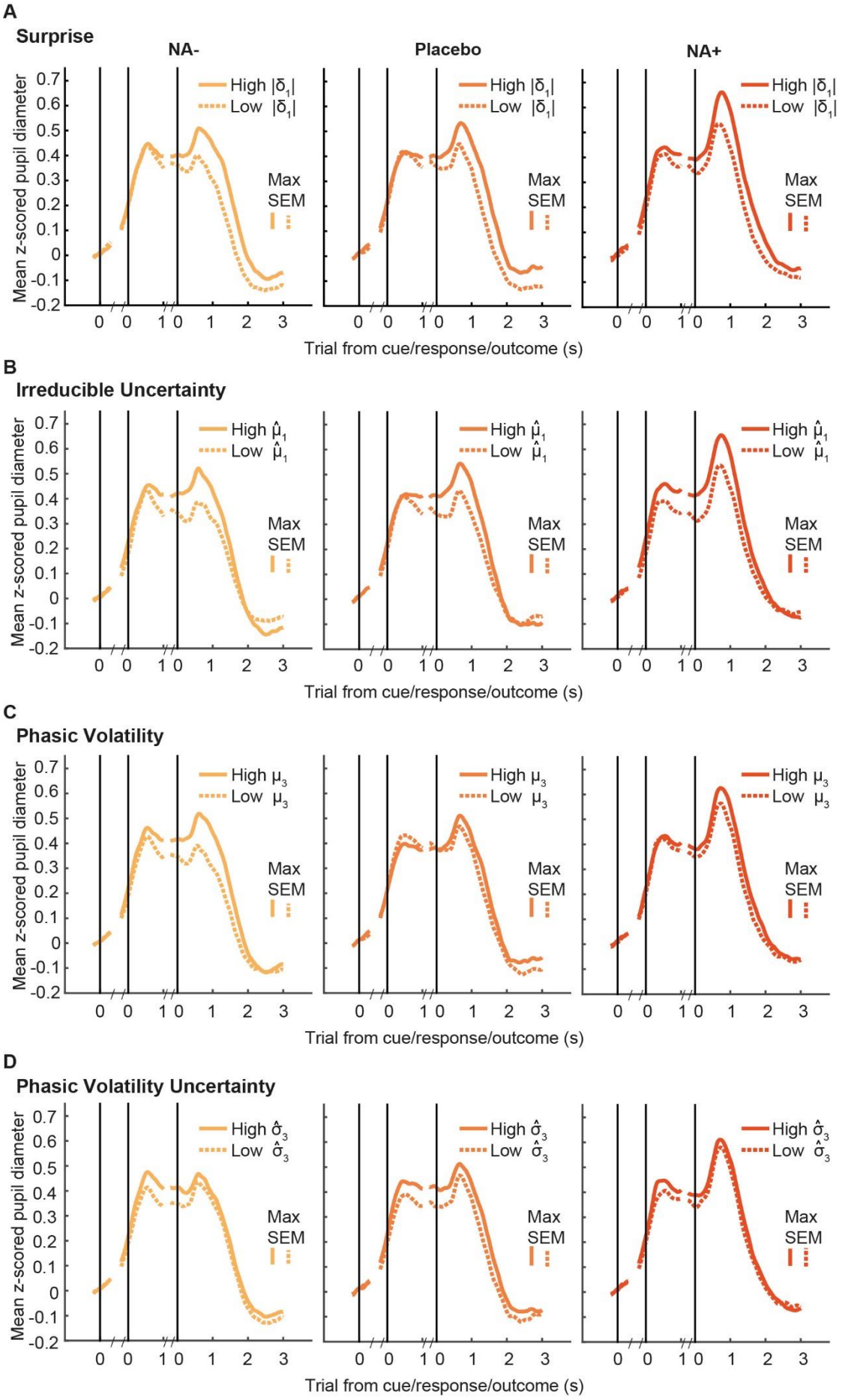
A
Surprise
NA-
Placebo
NA+
Mean z-scored pupil diameter
High |δ1|
Low |δ1|
Max
SEM
B
Irreducible Uncertainty
High μ̂1
Low μ̂1
C
Phasic Volatility
High μ3
Low μ3
D
Phasic Volatility Uncertainty
High σ̂3
Low σ̂3
Trial from cue/response/outcome (s)

Figure 6.11 Model-based analysis of pupil diameter. *Median splits indicated that, across drug-groups, high surprise ($|\delta_1|$; A) and high irreducible uncertainty ($\hat{\mu}_1$; B) were associated with an increased pupil diameter in the post-outcome period compared to low surprise and low irreducible uncertainty. High phasic volatility estimates (μ_3; C) were associated with increased post-outcome pupillary dilation in the NA- group only. There was a tendency for high phasic volatility uncertainty ($\hat{\sigma}_3$; D) to be associated with increased post-outcome pupil diameter, but the effect was not significant within individual drug-groups. There was also a tendency for high irreducible uncertainty and high volatility uncertainty to increase pupil diameter in the post-response period across drug-groups. Data are mean pupil diameter across participants. Error bars indicate the maximum SEM for each trial-type.*

6.4.3.5 Regression analyses: Placebo data

Regression analyses enabled the specific pupillary effects of participants' estimates of surprise, irreducible uncertainty, phasic volatility, and phasic volatility uncertainty during the post-response and post-outcome periods to be identified. Reassuringly, the regression analysis of the Placebo group pupil data identified a constant component of pupil diameter (Figure 6.12A) with a trial-wise trajectory strikingly similar to the mean pupil diameter trajectory shown in Figure 6.8. FDR-corrected t-tests indicated that pupil diameter was significantly greater than baseline during the 0-1s post-response period (all FDR-corrected $p<0.001$), and from 0-1.5s post-outcome (all $p\leq0.001$). Pupil diameter then decreased below baseline levels from 2.3-3.2s post-outcome (all $p\leq0.03$).

Surprise significantly increased pupil diameter from 1.2-2.3 and 2.7-3.2s post-outcome (all $p\leq0.04$); Figure 6.12B). Again, this result is reassuring given the positive effect of surprise and incorrect predictions on pupil diameter observed previously (Figure 6.11A).

Irreducible uncertainty had a tendency to increase pupil diameter 0-1s post-response (all uncorrected $p\leq0.03$), but this result did not survive correction for multiple comparisons. Irreducible uncertainty did significantly increase pupil diameter 0-1.s post-outcome (all FDR-corrected $p\leq0.04$; Figure 6.12C).

Phasic volatility estimates showed a tendency to decrease pupil diameter 0.8-1s post-response (uncorrected $p=0.038$), but this result did not survive correction for multiple comparisons (Figure 6.12D). No significant effects of phasic volatility uncertainty on pupil diameter were observed (Figure 6.12E).

In summary, pupil diameter tracked both surprise and irreducible uncertainty. Additional control regressors for cue type, response type and outcome type were included in the analysis but had no significant influence on pupil diameter (all FDR-corrected p>0.09).

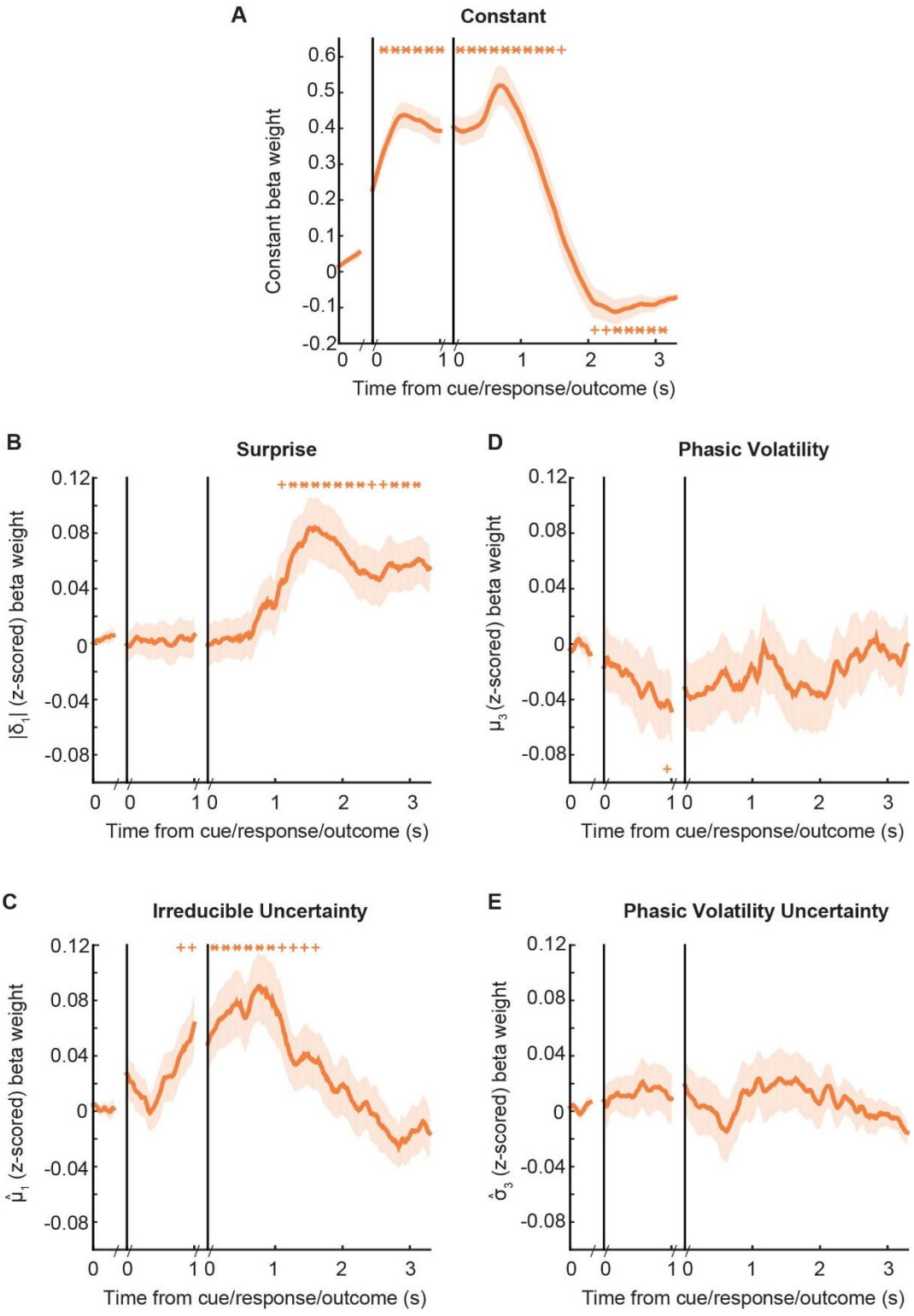

Figure 6.12 Regression analyses on Placebo pupil data. *The output indicates the effects of Placebo participants' beliefs on pupil diameter. The constant component of*

*pupil diameter strongly reflects the mean pupil diameter trajectory (see Figure 6.8). Data are mean ± SEM. * p<0.05 (FDR-corrected), $^{+}$ p<0.05 (uncorrected).*

6.4.3.6 Regression analyses: Noradrenergic manipulations

Next the output of the regression analyses for each of the active drug-groups was compared to the output of the regression analysis for the Placebo group (Figure 6.13). Compared to Placebo, individuals in the NA+ group showed a general tendency towards an increased pupil diameter in the post-outcome period, indicated by an augmented constant component of pupil diameter 0.7-1.3s and 1.7-2.3s post-outcome (all p≤0.04; Figure 6.13A). However, these results did not survive FDR-correction for multiple comparisons. While the NA- group showed a numerical decrease in the constant component of post-outcome pupil diameter compared to Placebo approximately 0.7-1.3s post-outcome, this decrease was not statistically significant (all uncorrected p≥0.12).

Compared to Placebo, surprise tended to have a reduced influence on pupil diameter in the NA- group 0-1s post-response (all uncorrected p≤0.05), and in the NA+ group 1.5-2.3s post-outcome (all uncorrected p≤0.05). However, these results did not survive correction for multiple comparisons (Figure 6.13B).

Irreducible uncertainty estimates had a significantly reduced effect on pupil diameter in the NA- group 0.5-1s post-response (all FDR-corrected p≤0.05; Figure 6.13C) and 0-1.8s post-outcome (all p≤0.05). In the NA+ group, a reduced effect of irreducible uncertainty on pupil diameter was only observed after outcome onset, specifically 0.7-1s and 1.3-1.7s post-outcome (all p≤0.05).

Compared to Placebo, phasic volatility estimates significantly increased pupil diameter in both drug-groups (Figure 6.13C). In both the NA- and NA+ groups, volatility increased pupil diameter 0.5-1s post-response (all FDR-corrected p≤0.02). For the post-outcome period, volatility increased pupil diameter 0-1.2s post-outcome in the NA- group, and 0-2s post-outcome in the NA+ group (all FDR-corrected p≤0.03).

Phasic volatility uncertainty estimates tended to decrease pupil diameter during the 0.2-0.5s post-response in the NA+ group compared to Placebo (all uncorrected p≤0.05; Figure 6.13D), but this result did not survive FDR-correction for multiple comparisons. There was no effect of NA- on pupil diameter modulation by phasic volatility uncertainty.

In summary, noradrenergic antagonism (NA-) modulated the pupil response to volatility and irreducible uncertainty. Boosting noradrenergic function (NA+) had the tendency to increase the general responsivity of the pupil, and also modulated the pupil response to volatility and irreducible uncertainty.

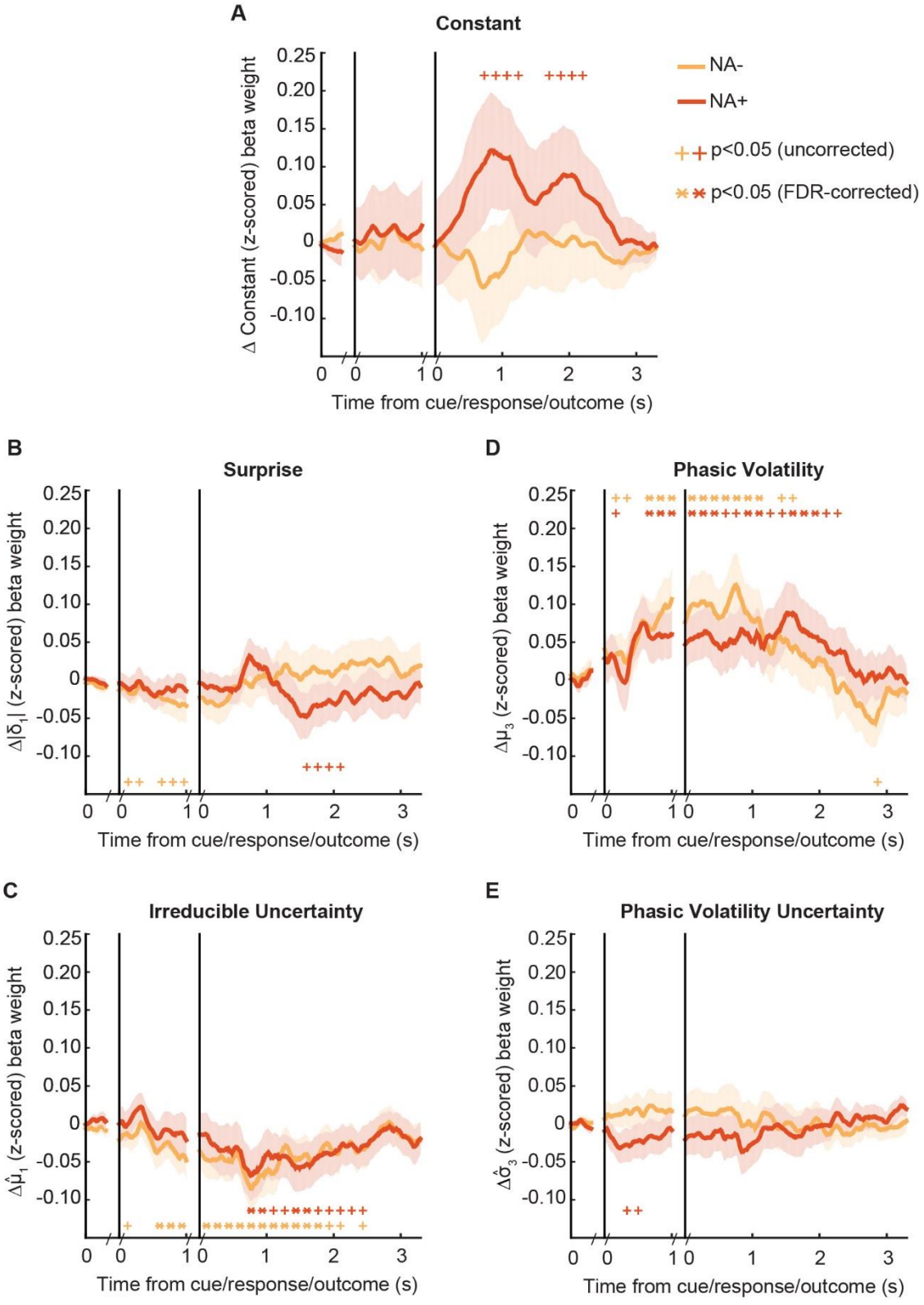

Figure 6.13 Regression analyses on NA- and NA+ pupil data. *The output indicates noradrenergic manipulations of participants' beliefs on pupil diameter. Data are Drug – mean Placebo, ± the standard error of the difference (SED). * p<0.05 (FDR-corrected), + p<0.05 (uncorrected).*

6.4.4 Learning rate

As in Chapter 4, the punctate change-points contained in the PLT's true generative process are detected implicitly as an increase in learning rate (α_1; Figure 6.14A). This implies that, as one would expect, higher learning rates following a true change-point reflect behaviour that is strongly controlled by recent outcomes.

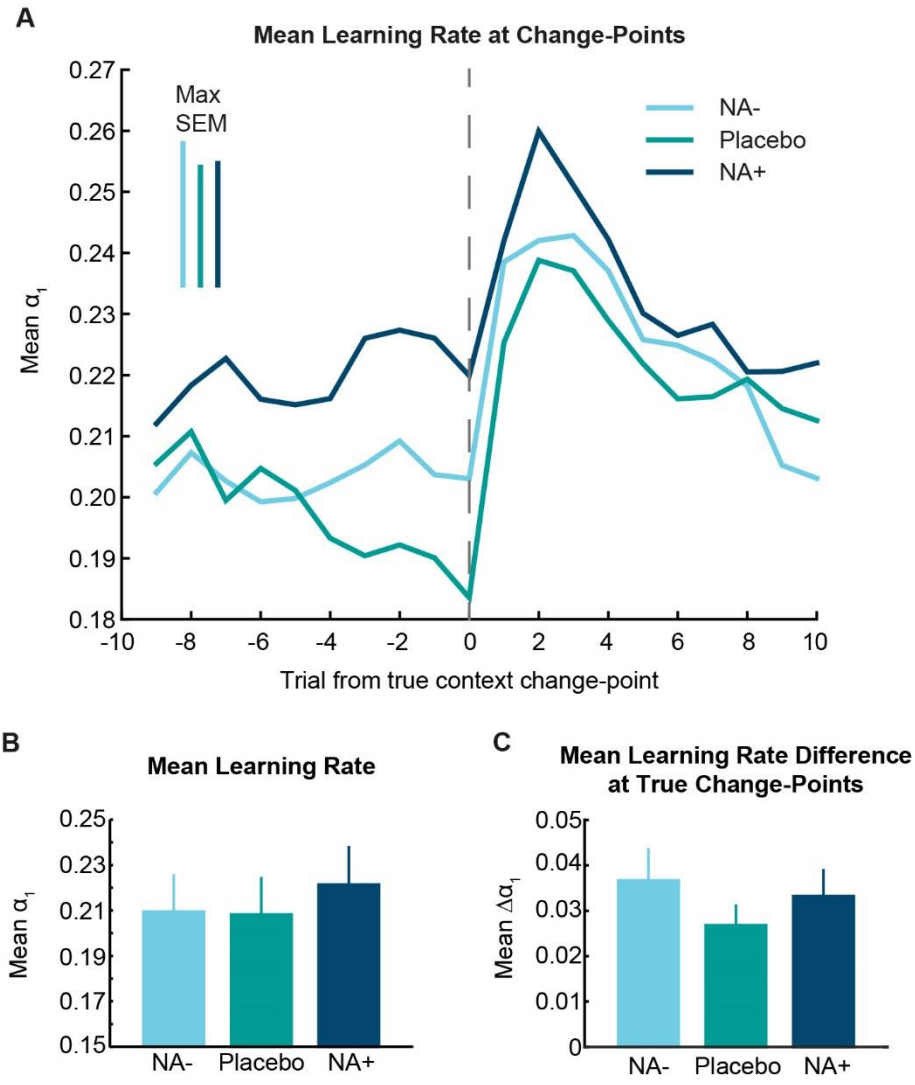

Figure 6.14 Mean learning rates across drug-groups. *(A) An increase in mean learning rate was identified following true context change-points. Error bars indicate the maximum SEM for each drug-group. (B) Mean learning rates across the PLT. (C) Mean difference in learning rate at all context change-points. For B and C, data are mean ± SEM.*

A one-way ANOVA demonstrated that mean learning rates across the PLT were equivalent across drug-groups ($F_{2,86}$=0.07, p=0.928; Figure 6.14B) and that there was no modulation by Δalertness ($F_{1,86}$=0.42, p=0.520). Moreover, the drug manipulations did

not alter the mean difference in learning rate across all types of context switch ($F_{2,86}$=0.61, p=0.547; Figure 6.14C). However, across the three drug-groups, there were differences in the degree to which learning rates changed across different types of context switch (Figure 6.15). Indeed, an 8 context switch-type x 3 drug RM-ANOVA indicated that the type of context switch significantly altered the difference in learning rate ($F_{2.80,25.15}$=3.99, p=0.021, η_p^2=0.31). There was a trend-level modulatory effect of drug-group ($F_{5.59,25.15}$=2.02, p=0.105, η_p^2=0.31, and no modulation by Δalertness (p=0.742).

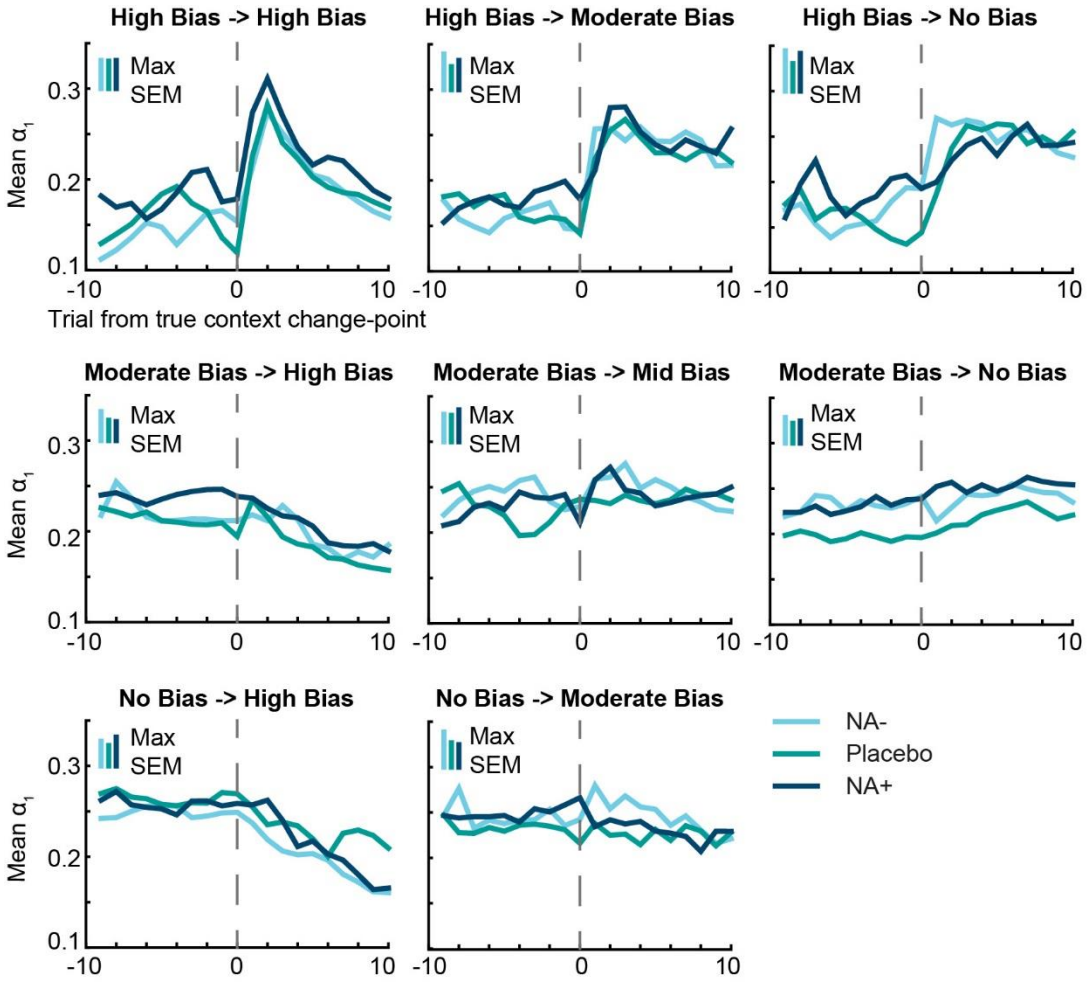

Figure 6.15 Mean learning rates at all types of context change-point. *Augmented increases in learning rate were observed for more obvious context switches (e.g., from a highly biased context to a context highly biased in the opposite direction, or from a highly biased to a moderately biased context) than for less obvious context switches (e.g., from a moderately biased to an unbiased context). Error bars indicate the maximum SEM for each drug-group.*

For exploratory purposes, I next assessed whether the drug manipulations altered the degree to which learning rates changed following the three most obvious context

changes (i.e., switches from a highly biased context to either a different highly biased, moderately biased or unbiased context) using three individual (FDR-corrected) one-way ANOVAs (Figure 6.16). No significant effect of drug was observed at switches to highly biased (p=0.974), moderately biased (p=0.792) or unbiased contexts (p=0.276). Note that the uncorrected comparisons were also non-significant (p=0.974, p=0.528 and p=0.092, respectively).

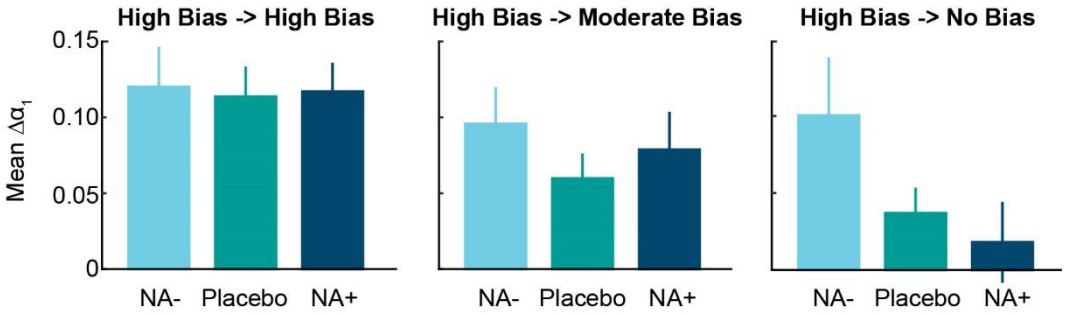

Figure 6.16 Mean learning rate differences at context change-points. *No significant effects of drug on the difference in learning rate were identified following a switch from a highly biased block to a different highly, moderately or unbiased block. Data are mean ± SEM, corrected for Δalertness.*

6.4.5 Behaviour vs pupil data

To ascertain whether the effect of participants' beliefs on pupil diameter was associated with behavioural changes during the PLT, correlational analyses were conducted on the volatility and irreducible uncertainty pupil beta weights, mean learning rates and mean performance scores.

The influence of volatility estimates on pupil diameter was negatively correlated with mean learning rate in the NA+ group (Pearson's r=-0.418, p=0.024; Figure 6.17A) but not in the NA- (r=0.021, p=0.911) or Placebo (r=-0.215, p=0.254) groups. Volatility's effect on pupil diameter was not correlated with mean performance score in any of the drug-groups (NA-: r=-0.218, p=0.247; Placebo: r=-0.156, p=0.411; NA+: r=-0.123, p=0.524; Figure 6.17B).

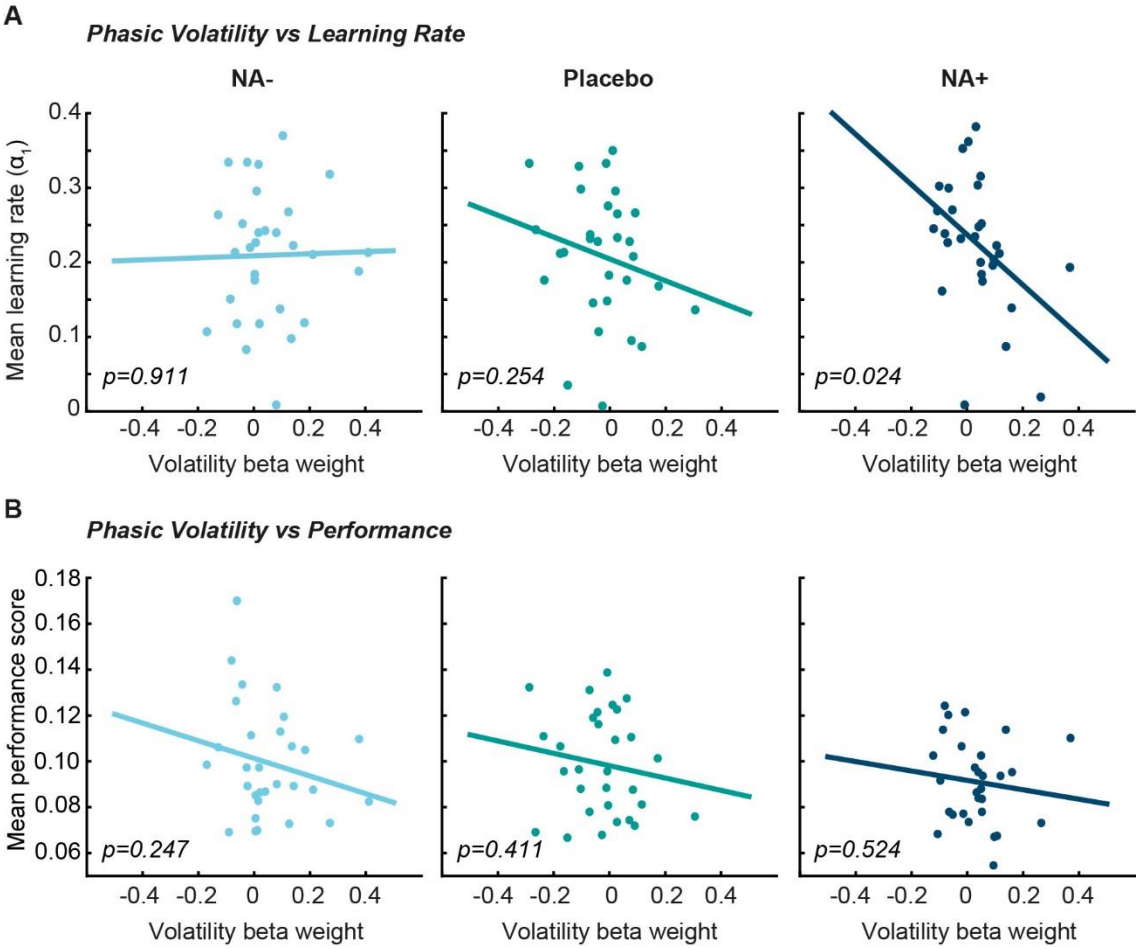

Figure 6.17 The effect of phasic volatility estimates on pupil diameter compared to behaviour. *The influence of phasic volatility estimates on pupil diameter was negatively correlated with mean learning rate in the NA+ group. No other significant correlations were identified.*

The influence of irreducible uncertainty estimates on pupil diameter were not correlated with mean learning rate in any of the drug-groups (NA-: r=-0.016, p=0.935; Placebo: r=-0.160, p=0.412; NA+: r=0.217, p=0.259; Figure 6.18A). Irreducible uncertainty's effect on pupil diameter was positively correlated with mean performance score in the NA+ group (r=0.0374, p=0.045; Figure 6.18B), but not in the NA- (r=0.106, p=0.579) or Placebo (r=0.006, p=0.976) groups.

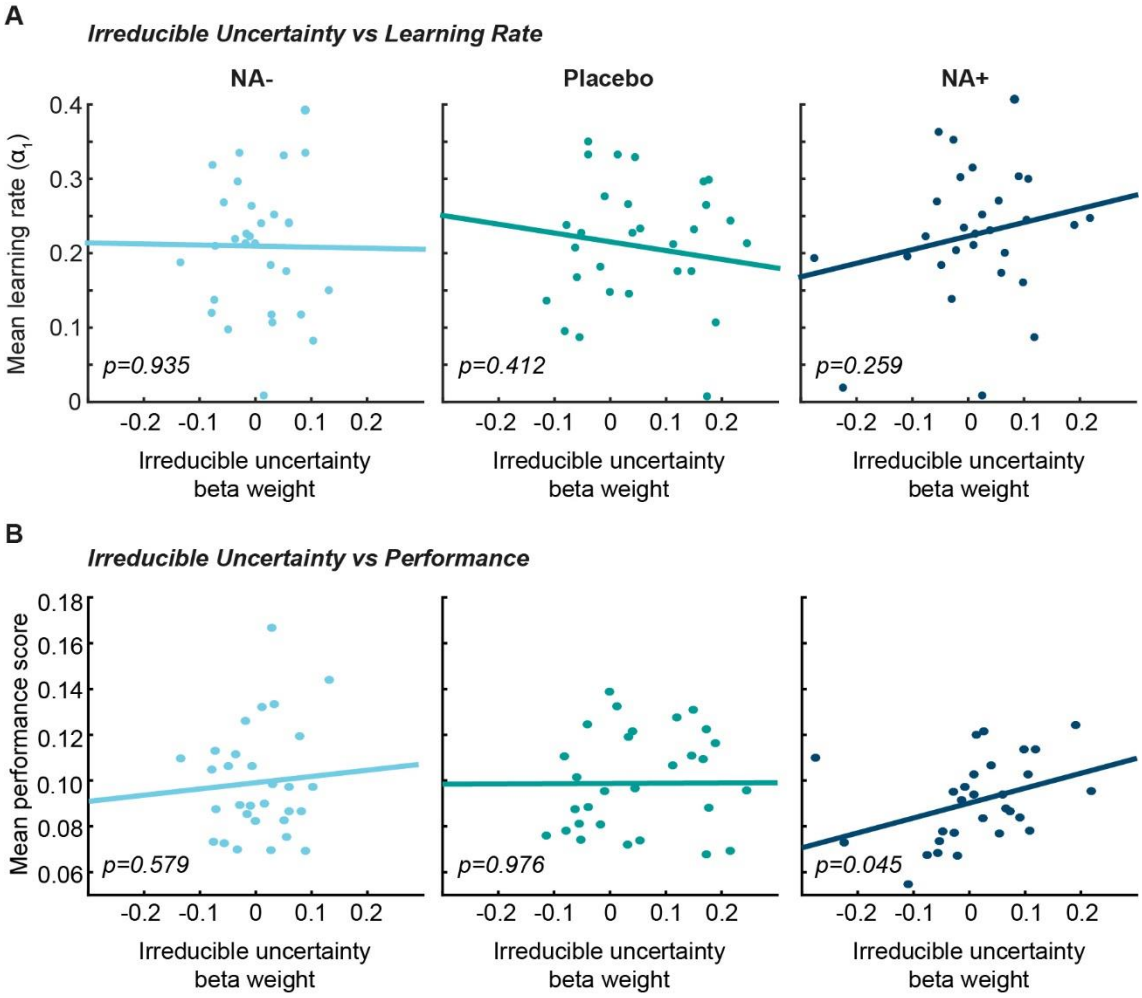

Figure 6.18 The effect of irreducible uncertainty on pupil diameter compared to behaviour. *The influence of irreducible uncertainty estimates on pupil diameter were positively correlated with mean performance score in the NA+ group. No other significant correlations were identified.*

6.4.6 Control analyses

6.4.6.1 Physiological and subjective control measures

Self-reported ratings for alertness and contentedness changed significantly over the course of the experiment ($F_{1.74,151.73}$=42.63, p<0.001, η_p^2=0.33; $F_{1.71,148.93}$=12.84, p<0.001, η_p^2=0.13 respectively); calmness ratings remained unchanged ($F_{2,174}$=0.88, p=0.42). Alertness ratings showed a significant time x drug interaction ($F_{3.48,151.73}$=3.34, p=0.016, η_p^2=0.07). A one-way ANOVA with drug as a between-subjects factor revealed that the degree to which alertness decreased between Baseline (Figure 6.1A) and the time corresponding to peak drug concentration (Post-Drug) varied between groups ($F_{2,87}$=6.56, p=0.002, η_p^2=0.13). More specifically, (FDR-corrected) pairwise

comparisons indicated that, compared to Placebo, the alertness-decrease was significantly more pronounced in the NA+ group (t_{58}=-3.59, p=0.002, *d*=-0.94).

Heart rate (HR) varied significantly with time ($F_{2,174}$=13.11, p<0.001, η_p^2=0.13) and this effect was modulated by drug-group ($F_{4,174}$=4.22, p=0.003, η_p^2=0.09). On average, all groups showed mild participant-specific HR decreases between Baseline and Post-Drug. The magnitude of HR deceleration differed between groups ($F_{2,87}$=5.13, p=0.008, η_p^2=0.11), with the deceleration being less pronounced in the NA- (t_{58}=2.24, p=0.027, *d*=0.59) and NA+ (t_{58}=3.10, p=0.006, *d*=0.81) groups compared to Placebo. Systolic blood pressure (BP) did not change significantly over time (p=0.783) and there was no modulation by drug-group (p=0.088). Diastolic BP did vary significantly with time ($F_{2,174}$=3.30, p=0.039, η_p^2=0.04) and this effect was modulated by drug-group ($F_{4,174}$=5.81, p<0.001, η_p^2=0.12). Indeed, in line with the known effect of NA on BP, an effect of drug-group on the participant-specific difference in diastolic BP between Baseline and Post-Drug ($F_{2,87}$=13.42, p<0.001, η_p^2=0.24) was driven by a significant increase in BP in the NA+ group compared to Placebo (t_{58}=3.22, p=0.004, *d*=0.85) and a trend-wise decrease in BP in the NA- group (t_{58}=-1.90, p=0.061, *d*=-0.50). A summary table of the subjective and physiological measures is reported in (Table 6.3).

6.4.6.2 Parameter correlations

The parameters entered into the pupil regression analyses were not highly correlated (Figure 6.19). The highest correlation existed between surprise ($|\delta_1|$) and phasic volatility estimates (μ_3) (mean for NA-: r=0.534; Placebo: r=0.536; NA+: r=0.498). The absolute mean for all other correlations was r≤0.055 for NA-, r≤0.134 for Placebo, and r≤0.056 for NA+.

		NA-	Placebo	NA+
Alert-ness	Baseline	70.9 ± 2.8	68.9 ± 2.9	69.3 ± 2.8
	Post-Drug	61.8 ± 3.5	64.7 ± 3.0	52.2 ± 3.6
	Post-Task	56.6 ± 3.2	60.4 ± 2.9	49.9 ± 3.8
Calm-ness	Baseline	72.2 ± 3.2	67.6 ± 2.7	61.5 ± 2.8
	Post-Drug	67.2 ± 4.2	72.4 ± 2.8	64.9 ± 3.0
	Post-Task	64.4 ± 3.8	67.0 ± 2.9	66.0 ± 2.6
Content-edness	Baseline	72.9 ± 2.6	78.0 ± 2.5	71.4 ± 2.5
	Post-Drug	69.4 ± 2.6	75.5 ± 2.2	64.6 ± 2.9
	Post-Task	65.7 ± 2.4	73.2 ± 2.6	62.1 ± 3.9
HR	Baseline	73.9 ± 2.1	71.8 ± 1.6	69.2 ± 2.3
	Post-Drug	71.1 ± 2.1	64.1 ± 1.6	68.2 ± 2.2
	Post-Task	68.3 ± 1.8	64.1 ± 1.6	69.3 ± 2.2
Systolic BP	Baseline	121.6 ± 2.8	118.6 ± 2.4	120.1 ± 2.6
	Post-Drug	117.5 ± 2.6	116.7 ± 2.4	124.2 ± 2.9
	Post-Task	120.3 ± 2.7	118.7 ± 2.8	121.9 ± 2.3
Diastolic BP	Baseline	75.2 ± 1.6	74.3 ± 1.5	74.1 ± 1.6
	Post-Drug	71.4 ± 1.5	74.1 ± 1.5	80.1 ± 2.0
	Post-Task	73.5 ± 2.4	77.3 ± 1.5	79.4 ± 2.2

Table 6.3 Subjective and physiological measures for each experimental group. *Readings were taken at baseline, immediately before participants started the PLT (i.e., when the drugs were at their most active; Post-Drug), and after completing the PLT and CT (Post-Task). Data are mean ± SEM.*

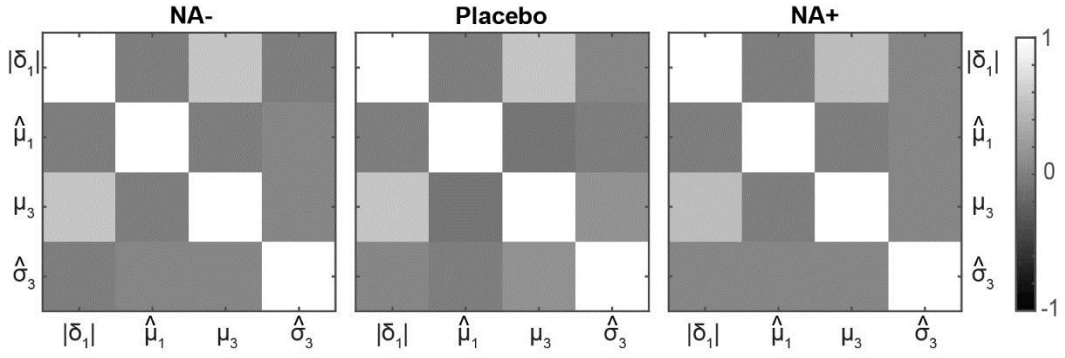

Figure 6.19 Model parameter correlations. *The parameters used in the regression analyses were not highly correlated with each other (all mean absolute r≤0.534 for NA-, r≤0.536 for Placebo, and r≤0.498 for NA+).*

6.4.6.3 Decision times during the PLT and CT

On average, mean and minimum decision times during the CT tended to be faster than during the PLT across drug-groups (Table 6.4). This is to be expected given that participants knew with certainty which cue and which outcome would be presented on each trial.

		NA-	Placebo	NA+
PLT	**Mean Decision time (ms)**	651.6 ± 20.7	661.7 ± 22.0	676.5 ± 20.5
	Mean Minimum Decision time (ms)	300.3 ± 14.9	315.1 ± 14.2	327.6 ± 14.1
CT	**Mean Decision time (ms)**	392.3 ± 21.2	422.1 ± 23.4	428.8 ± 21.4
	Mean Minimum Decision time (ms)	232.0 ± 12.2	234.8 ± 11.6	242.0 ± 14.0

Table 6.4 Mean and minimum decision times during the PLT and CT. *Data are mean ± SEM.*

6.5 Discussion

By combining a probabilistic learning task with pharmacological manipulations, pupillometry and a hierarchical Bayesian learning model, it was possible to characterise the impact of NA and dynamic uncertainty computations on changes in pupil diameter. Implementing a unified framework of hierarchically-related forms of uncertainty meant that I could assess the degree to which individuals' computations of irreducible uncertainty, surprise, phasic volatility and phasic volatility uncertainty were reflected by the pupil, and compare the effects of two drugs, with different effects on NA neurotransmission, on individuals' subjective beliefs, behaviour and pupillary responses.

6.5.1 Baseline pupil diameter reflects individual irreducible uncertainty

The finding that baseline pupil diameter increases with increasing irreducible uncertainty replicates our previous finding with a visual version of the PLT (de Berker et al., 2016). It also echoes a finding by Nassar et al. that baseline pupil diameter reflects the reliability with which recent event history indicates the current probabilistic relationship between environmental events during predictive inference (Nassar et al., 2012). Strikingly, the inverted-U relationship observed between pupil diameter and irreducible uncertainty in

the present study was strongly reflected in the relationship between irreducible uncertainty and decision time, indicative of a tight link between irreducible uncertainty computations, baseline pupil diameter and behaviour. However, these relationships do not appear to be dependent on NA. Indeed, neither downregulating NA neurotransmission with prazosin nor upregulating NA neurotransmission with reboxetine appeared to alter the relationships between irreducible uncertainty, baseline pupil diameter and decision time.

6.5.2 Pupil diameter tracks irreducible uncertainty and surprise

Using median splits to compare peak pupil diameter on trials high in subjective surprise and irreducible uncertainty suggested that each of the perceptual quantities had a positive effect on pupil diameter in the Placebo group, replicating our previous work (de Berker et al., 2016). Estimates of phasic volatility and phasic volatility uncertainty did not appear to modulate pupil diameter. However, this method of analysis only offers a relatively crude way of repeatedly categorising trials according to subjective beliefs.

Importantly, in the present study, regression analyses were implemented to pinpoint the relative contributions of individuals' subjective beliefs on pupil diameter during the post-response and post-outcome periods. The Placebo group showed increases in pupil diameter during the post-outcome period that were dependent on subjective estimates of irreducible uncertainty and surprise, echoing the median split analyses used both here and in previous work (de Berker et al., 2016). Pupillary dilation due to post-outcome surprise has also been observed by other research groups (Preuschoff et al., 2011; Browning et al., 2015), albeit with subtly different quantifications of surprise. In particular, my finding of a positive effect of surprise on pupil diameter 1.2-3.2s post-outcome replicates Browning et al.'s observation that surprise increases pupil diameter in a 1-3s post-outcome window during probabilistic learning (Browning et al., 2015; c.f. Figure 6.12B and Figure 6.20A).

As anticipated by the median split method, regression analyses did not reveal a significant effect of individuals' estimates of phasic volatility or phasic volatility uncertainty on pupil diameter. The lack of an effect of volatility estimates on pupil dilation contradicts a finding by Browning et al. that volatility increases post-outcome pupil diameter (Browning et al., 2015; c.f. Figure 6.12C and Figure 6.20B). However, the authors observed the pupil diameter increase 2-5s post-outcome, a window that lay outside the limits of the minimum ITI used in the current task. Therefore, it is possible that the post-outcome period used in the present experiment was not long enough to capture this effect.

Figure 6.20 The effects of surprise and volatility on post-outcome pupil diameter during aversive probabilistic learning observed by Browning et al. *(A) The post-outcome pupillary response to surprise is strikingly similar to that observed in the present study (c.f. Figure 6.12B). (B) Browning et al. also identified a positive effect of environmental volatility on pupil diameter post-outcome, which I did not replicate (c.f. Figure 6.12C). Figure adapted from Browning et al., 2015.*

Another notable difference between Browning et al.'s study and the experiment conducted here is the behavioural paradigm used to generate volatility. Browning et al. manipulated the volatility of trials in a block-wise manner so that participants had to learn one probabilistic relationship during a stable block of 90 trials (i.e., approximately three times the block-length used in the present study) or track multiple changes in a probabilistic relationship during a volatile block of 90 trials. This method may have offered a superior means by which to manipulate individuals' volatility estimates and may thus have produced larger effects on pupil diameter. However, since Browning et al.'s outcome stimuli were electric shocks rather than auditory stimuli, the possibility that the effect of volatility estimates on pupil diameter was increased under exposure to aversive stimuli cannot be ruled out. Indeed, our previous work has highlighted the relevance of pupillary dilation as an acute stress response during learning under uncertain threat of aversive stimuli (de Berker et al., 2016). Nonetheless, it would certainly be interesting to adopt Browning et al.'s method of manipulating the volatility of a probabilistic relationship between auditory cues and auditory outcomes. As such, it would be possible to determine whether this paradigm has a superior ability to capture a post-outcome effect of volatility estimates on pupil diameter in the absence of aversive stimuli.

6.5.3 Noradrenaline influences pupillary responses to trial outcome, volatility and irreducible uncertainty

6.5.3.1 Noradrenaline modulates the constant component of pupil diameter

The opposing noradrenergic manipulations tended to have opposite effects on the constant component of pupil diameter. Upregulating NA neurotransmission with reboxetine resulted in a trend-wise increase in post-outcome pupil diameter, whereas NA antagonism under prazosin tended to decrease the post-outcome pupil diameter. While these results should be interpreted with caution since they did not survive correction for multiple comparisons, they do fit well with previous literature that has assessed the impact of noradrenergic agents on pupillary dynamics. Indeed, α2-adrenoceptor agonists, such as clonidine, which decease the activity of central noradrenergic neurons (and thus have a similar effect to α1-adrenoceptor antagonists such as prazosin), have been shown to decrease baseline pupil diameter and increase spontaneous pupillary fluctuations. In contrast, α2-adrenoceptor antagonists, such as yohimbine, which have the opposite effect on central NA, have been shown to have the opposite effect on pupils, i.e., an increase in baseline pupil diameter and decreased pupillary fluctuations (Phillips et al., 2000b). Moreover, the fact that these noradrenergic effects on the constant component of pupil diameter are observed post-outcome, but not post-cue or post-response, suggests that they are event-specific.

6.5.3.2 Noradrenaline modulates pupillary responses to volatility

Both noradrenergic manipulations increased the effect of subjective volatility on post-response and post-outcome pupil dilation. This interaction between NA, volatility and pupil diameter sits well with the finding in Chapter 4 that NA plays an important role in learning under uncertainty that arises from environmental volatility (Yu and Dayan, 2005; Payzan-LeNestour et al., 2013; Marshall et al., 2016) and with suggestions that pupil diameter offers an indirect measure of noradrenergic neuronal activity in the LC (Aston-Jones and Cohen, 2005a; Varazzani et al., 2015; Joshi et al., 2016).

Nonetheless, the fact that the direction of the effect on pupil diameter was identical under two drugs with supposedly opposing effects on NA neurotransmission is puzzling. This likely speaks to the complexities and subtleties of the neuromodulatory systems with which the pharmacological agents interacted. Prazosin acts as an antagonist at α1-adrenoceptors, a high density of which exist in the LC, at least in rats (Jones et al., 1985; Stone et al., 2004), and in the human neocortex (Zilles et al., 1993). As such, one would expect it to have reduced LC-NA neurotransmission to the cortex. In contrast, reboxetine

is a selective NA reuptake inhibitor (SNRI) that works by blocking the action of the NA transporter (NET), thereby slowing the rate at which NA is cleared from the synaptic cleft. Since reboxetine increases extracellular NA concentrations, it is thought to increase NA neurotransmission.

However, it has been proposed that, while SNRIs primarily increase NA levels due to reuptake inhibition, secondary indirect activation of inhibitory presynaptic α2-autoreceptors reduces noradrenergic activity in areas such as the LC (Invernizzi and Garattini, 2004). It is therefore possible that reboxetine may have actually reduced noradrenergic firing in the LC in some individuals. However, given the observation of a trend-wise increase in the constant component of pupil diameter under reboxetine, in line with previous accounts that upregulating NA results in pupillary dilation (Phillips et al., 2000b), this seems unlikely in the present study, at least on average. It has also been proposed that the net effect of a SNRI's two actions likely depends on dosage and on an individual's baseline NA activity (Coull, 2001; de Rover et al., 2012). Future studies will need to determine whether any polymorphisms in the *NET* gene are associated with inter-individual differences in baseline NA activity and, if so, whether these polymorphisms are linked to altered responses to SNRIs and altered pupillary responses to volatility estimates.

Furthermore, the extent to which pharmacological manipulations of NA modulate phasic and tonic modes of NA activity is currently unclear. As discussed in Chapter 4, the neurophysiological literature has described two functional modes of NA neurotransmission in the LC (Aston-Jones and Cohen, 2005b; Bouret and Sara, 2005). A phasic mode, characterised by a relatively low baseline firing rate and high phasic responsiveness to task relevant stimuli, has been linked to enhanced task engagement. A tonic mode (lacking phasic activity) has been linked to increased distractibility, attention-shifting and exploratory behaviour (Aston-Jones et al., 1994; Usher et al., 1999; Aston-Jones and Cohen, 2005b, but see Jepma et al., 2010). Some pharmacological manipulations of NA have been shown to shift the balance between phasic and tonic noradrenergic activity.

In Chapter 4, I discussed the possibility that prazosin may have enabled more phasic NA responsiveness to emerge under suppression of tonic NA firing. Related to this point, a recent study in rats found that the NET-blocker atomoxetine reduces baseline LC activity while preserving the stimulus-evoked phasic LC response, leading to an increase in phasic relative to tonic LC activity (Bari and Aston-Jones, 2013). It is therefore possible that drugs that block NET, including reboxetine, enhance neural responses to stimuli that

evoke large LC responses. However, it should be noted that atomoxetine also blocks the serotonin transporter (SERT) and dopamine transporter (DAT), meaning that its effects on phasic and tonic NA activity may not be NET-specific. Further investigations utilising electrophysiological recordings in animals are required to characterise the relative impact of other noradrenergic drugs on phasic and tonic NA activity.

In summary, the fact that the two noradrenergic manipulations used in the present study modulated the pupillary response to volatility provides weight to the notion that pupil diameter reflects, in part, noradrenergic volatility computations. However, to determine the precise neurophysiological bases of these responses, future work and continued cross-talk between human and animal research are required to pinpoint the effects of different pharmacological manipulations on NA neurotransmission.

6.5.3.3 Noradrenaline modulates pupillary responses to irreducible uncertainty

Both noradrenergic manipulations decreased the effect of subjective irreducible uncertainty on post-outcome pupil dilation. NA antagonism under prazosin also decreased the post-response effect of irreducible uncertainty on pupil diameter. Again, the fact that the effect was negative in both drug groups is surprising, but the same considerations regarding the complex effects of pharmacological manipulations on intricate neuromodulatory systems discussed in the previous section apply here.

Importantly, the finding of an interaction between NA, irreducible uncertainty and pupil diameter highlights the need to consider the impact of multiple parameters when using pupil diameter as a proxy for subjective uncertainty computations. As is apparent in the present experiment, the pupil does not reflect a single belief but is rather modulated by estimates of irreducible uncertainty, surprise and volatility. Thus, unified frameworks of uncertainty, such as that offered by the HGF, are an important tool for pinpointing the relative contributions of individuals' subjective estimates on pupil diameter.

6.5.4 Pupil dilation follows a motor response indicating a decision

In addition to the post-outcome pupillary responses observed in the present study, pupil dilation also occurred following the button-press response participants made to indicate their prediction as to the trial's outcome. Since post-response pupillary dilation occurred during both the PLT, where there was always a degree of uncertainty about whether the decision was correct, and during the CT, where participants were certain of the outcome that would follow each cue, it appears that this pupillary response is, at least in part, driven by the motor response used to report the choice. Moreover, the fact that post-response pupillary dilation was augmented and sustained under conditions of uncertainty

during the PLT, compared to certainty during the CT, is indicative of an additional dilatory effect of post-decisional uncertainty about whether the predicted outcome was correct or incorrect. Indeed, in line with this, the regression analysis on the Placebo pupil data indicated a tendency for irreducible uncertainty to increase pupil diameter during the post-response period.

In accordance with this interpretation, it has been previously proposed that pupil-linked neuromodulatory systems are activated by the termination of decision processes and consequently that these systems affect the post-decisional brain state. Indeed, pupil dilation has been linked to the final choice terminating a decision process (Einhäuser et al., 2008; Hupé et al., 2009; Einhauser et al., 2010). Some research groups have suggested that decision-related noradrenergic brainstem activity, and thus pupil dilation, is driven by an individual's final commitment to a choice (Aston-Jones and Cohen, 2005a; Einhäuser et al., 2008), and others that it is driven by the motor response used to report that choice (Hupé et al., 2009). The present results are compatible with the notion that post-response pupillary dilation is modulated by both factors.

In contrast, de Gee et al. have argued that pupil dilation is actually primarily driven during the course of making a decision, rather than once a decision has been made (de Gee et al., 2014). Specifically, sustained pupil dilation has been shown to coincide with the course of decision formation during an isoluminant visual detection task in which participants were required to decide whether or not a visual contrast signal was embedded in dynamic noise. Moreover, the magnitude of this intra-decisional pupil dilation was found to be greater than the transient increase in pupil diameter that occurred following a decision indicated by button-press. The noradrenergic intra-decisional computations of uncertainty that this pupillary dilation may reflect are notionally sensible since estimating uncertainty before choice commitment would allow for anticipatory behavioural adaptation. Indeed, Urai et al. recently demonstrated that pupil dilation occurring after a perceptual choice but before feedback not only reflects decision uncertainty (i.e., the probability that a choice was correct given the sensory evidence) but also predicts subsequent behavioural biases (Urai et al., 2017).

Unfortunately, it is not possible to determine with confidence whether decision-making during the PLT used in the present experiment had a dilatory effect on pupils, or whether any dilatory effect was modulated NA, because the importance of decision speed was stressed to participants, meaning the time period between cue and response was often short. However, pupil diameter did appear to start increasing from cue-onset (Figure 6.8), suggesting a potential influence of an intra-decisional component on pupil diameter.

Nevertheless, since pupil diameter also tends to increase from cue-onset during the CT (Figure 6.9), this component does not seem specific to uncertainty modulations.

It should also be noted that, in contrast to de Gee's study, the design of the present task meant that irreducible uncertainty could be computed on the basis of trial history, and therefore that participants could estimate its current magnitude before a new trial began. Indeed, this was reflected by an increase in baseline pupil diameter with increasing irreducible uncertainty estimates (Figure 6.7). As such, an intra-decisional increase in pupil diameter, possibly reflecting an uncertainty computation, would not necessarily be comparable in this instance. In sum, while a potential additional influence of intra-decisional processes on pupil diameter cannot be ruled out, the present results do suggest that the act of making a motor response to indicate a decision does have a dilatory effect on pupils.

6.5.5 The link between behaviour and pupillary responses to subjective beliefs is unclear

The effects of the noradrenergic manipulations on behaviour during the PLT were subtle, with boosted NA neurotransmission under reboxetine leading to a smaller improvement in (model-agnostic) task performance across the course of biased blocks compared to Placebo. There was no difference in (model-based) learning rates across drug-groups.

Any associations between the effect of participants' beliefs on pupil diameter and behavioural changes during the PLT were also subtle. Indeed, correlations between pupillary responses and behaviour only existed in the NA+ group. Here, the influence of volatility and irreducible uncertainty estimates on post-outcome pupil diameter were negatively correlated with mean learning rate, and positively correlated with mean task performance score, respectively.

As such, how pupillary responses to subjective beliefs are linked to behaviour is unclear. This is likely to be, at least in part, due to the fact that performance was very good across all three drug-groups, with very little behavioural difference under the noradrenergic manipulations. Jepma et al. have recently offered alternative evidence that atomoxetine, which upregulates NA, modulates learning rate following an environmental switch. Critically, the direction of this effect was found to be dependent on an individual's baseline learning rate under placebo (Jepma et al., 2016). Due to the between-subjects design of the current experiment, it is not possible to replicate this analysis. Nonetheless, it would be fruitful to repeat the present experiment with a within-subjects placebo session to determine whether baseline-corrected learning rates correlate with the

pupillary response to volatility estimates, and thus whether there is any association between pupillary and behavioural responses to volatility.

6.5.6 A unified computational framework of uncertainty will facilitate future comparisons between different drugs and neuromodulatory systems

While the notion that pupil diameter offers an indirect measure of noradrenergic neural activity in the LC has received particular attention (Aston-Jones and Cohen, 2005a; Varazzani et al., 2015; Joshi et al., 2016), activity in the cholinergic basal forebrain has also been suggested to modulate pupil dilation (Yu, 2012). Pupil dilation is controlled by two muscles: the sphincter for contraction and the dilator for dilation. These two muscles are innervated by ACh and NA respectively. Both the muscarinic ACh antagonist scopolamine and the nicotinic ACh antagonist mecamylamine have been shown to increase pupil diameter in healthy elderly individuals (Little et al., 1998). Moreover, Alzheimer's disease patients, who have a severe and specific cholinergic deficit, have been found to have larger pupil diameters than healthy control individuals, both tonically and in reflexive response to light. Administering these patients an acetylcholinesterase inhibitor, which increases extracellular ACh concentrations, leads to pupillary responses reminiscent of those in healthy individuals (Fotiou et al., 2009). Furthermore, it has recently been demonstrated that pupillary fluctuations are highly correlated with activity of both NA and ACh projections to the cortex (Reimer et al., 2016).

Further work is needed to determine whether cognitive control of pupil diameter is under both ACh and NA influence, as opposed to NA alone, and to elucidate how these two neuromodulatory systems interact. In particular, there is a need for continued cross-talk between animal and human research so that electrophysiological recordings of neuromodulatory neurons can be linked to pupillary fluctuations in animals (Joshi et al., 2016) and to the effects of different pharmacological manipulations on behaviour and pupil dilation in humans. A unified computational framework of uncertainty, such as that applied in the present study, offers an ideal tool with which to make comparisons between the effects of different drugs and different neuromodulatory systems on dynamic uncertainty computations and their impact on pupil diameter.

It should also be noted that NET, the transporter targeted by reboxetine, can also reuptake extracellular dopamine (DA) (Bymaster et al., 2002; Swanson et al., 2006; Koda et al., 2010). Due to the extensive connectivity of its neuromodulatory network, DA, like NA, holds the potential to modulate synaptic transmission across the brain. Moreover, LC activity can trigger co-release of DA from noradrenergic terminals, at least in rats (Devoto and Flore, 2006), and there are bidirectional projections between the

dopaminergic ventral tegmental area and the noradrenergic LC (Sara, 2009). The neurocognitive literature has established that DA neurons respond to novel and unexpected stimuli, and that individuals with Parkinson's disease, and thus DA dysfunction, have an impaired capacity to switch between task-specific behaviours (Cools et al., 2001a; Cools and Robbins, 2004; Wise, 2004). As such, DA appears well-placed to support behavioural switching, which is an adaptive process under environmental volatility. Furthermore, in Chapter 4, DA was found to sensitise motor responses to subjective estimates of phasic volatility. Therefore, at present, additional dopaminergic modulation of the pupillary response to volatility cannot be ruled out. Again, pharmacological manipulations of the dopamine network within the same behavioural and computational framework would help to address this.

6.5.7 Conclusion

In summary, the results presented in this chapter offer novel insight into the relative contributions of uncertainty, surprise, volatility and noradrenaline to dynamic changes in behaviour and pupil diameter. Baseline pupil diameter strongly reflects an individual's belief about the current relationship between environmental events, as reflected by their irreducible uncertainty estimates. Dynamic pupillary dilation tracks both subjective irreducible uncertainty and surprise. NA modulates pupillary responses to irreducible uncertainty and volatility estimates. Pupillometry may therefore offer a useful proxy for computations of uncertainty, surprise and volatility, which appear to be, at least in part, dependent on NA. Importantly, changes in pupil diameter reflect dynamic beliefs about several perceptual parameters. This means that, while pupillometry offers a cheap and simple adjunct to behavioural paradigms, it should be used with suitable caution. Indeed, unified computational frameworks of uncertainty, such as the HGF, are required to fully capture the relative contributions of uncertainty, surprise and volatility to pupillary dilation. Future work utilising consistent behavioural and computational frameworks to contrast the impact of different pharmacological agents will help to elucidate the complexities of noradrenergic (and possible cholinergic and dopaminergic) modulations of behaviour and pupil diameter during learning.

7 General discussion

Learning the world's underlying statistical structure enables individuals to predict the likelihood of future environmental events, facilitating anticipatory action preparation and the execution of fast, accurate motor responses. In this thesis I have shed light upon the neuromodulatory processes that contribute to learning and response modulation in dynamic, probabilistic environments by examining the impact of pharmacological manipulations of noradrenaline (NA), acetylcholine (ACh) and dopamine (DA) (**Chapters 4** and **6**), and genetic variations in DA neurotransmission (**Chapter 5**), within a unified computational framework of uncertainty (**Chapter 3**). I have demonstrated that NA and ACh modulate learning under uncertainty. Specifically, NA influences learning of uncertain events that arise due to the environment's volatility. Further, dynamic noradrenergic computations of uncertainty and volatility can be measured indirectly using pupillometry. ACh balances the attribution of uncertainty to chance fluctuations within environmental contexts or to gross environmental violations following a contextual switch. In contrast, DA supports the use of perceptual estimates, namely volatility representations, to engender adaptive motor responses. Since each experimental chapter contains a relatively extensive discussion of the issues pertinent to that study, in this summary I draw together some common threads between the experiments, discuss the implications and limitations of this body of work, and formulate suggestions for future extensions to the field.

7.1 Benefits and limitations of the behavioural paradigms

7.1.1 The probabilistic serial reaction time task

A key benefit of the novel probabilistic serial reaction time task (PSRTT) implemented in **Chapters 4** and **5**, is the scope to characterise learning *and* response modulation under irreducible, estimation and volatility uncertainty. Its design created a more complex, and arguably more ecologically valid, scenario than earlier paradigms that explicitly signalled contextual rules and switches (Galea et al., 2012; Bestmann et al., 2014). Instead, individuals were required to infer a current environmental context for themselves and adapt to contextual changes. Within the framework of a novel instantiation of the Hierarchical Gaussian Filter (HGF) model, it was possible to interrogate the relative contributions of NA, ACh and DA to learning of the task's contextual probabilistic rules and to motor response modulation in light of individuals' beliefs about those rules (Marshall et al., 2016).

However, despite offering a paradigmatic advance, one should be mindful of the fact that the PSRTT is not fully representative of real-world learning and response modulation. Indeed, future work employing alternative task designs will be needed to verify the generality of the effects reported in **Chapter 4** to alternative behavioural paradigms with and without learning, reward, prediction and action. For instance, while using a paradigm that combined learning and action has the aforementioned merits, one could argue that indirectly inferring participants' predictions from their reaction times (RTs) is not the cleanest approach for investigating learning under different sources of environmental uncertainty. Indeed, contrasting the impacts of pharmacological manipulations of NA and ACh within a volatile predictive learning paradigm would offer a useful means by which to validate my finding that NA and ACh play respective roles in mediating learning between and within environmental contexts, in the absence of motor response modulation. Moreover, by introducing reward to the PSRTT, it might be possible to identify dopaminergic contributions to motor response modulation under reward prediction error (PE) rather than sensory PE.

In addition, while the use of speeded button-press responses and a novel response model enabled me to quantify the modulation of an individual's RTs by their perceptual beliefs, the approach offers little insight into how uncertainty modulates the quality of executed actions. The predominant approach to studying behavioural adaptation to perceptual estimates has been to examine when individuals initiate an action (Beierholm et al., 2013; Guitart-Masip et al., 2014; Vossel et al., 2014a, 2014b, 2015; Marshall et al., 2016). However, this only captures part of the picture. Environmental sources of uncertainty also influence how individuals execute actions (Bays and Wolpert, 2007). A simple yet fruitful extension to the work in this thesis would be to have individuals use a force button-box to complete the PSRTT. By requiring individuals to respond to the presentation of trial-wise stimuli by pressing an appropriate button quickly and within a required force range, it would be possible to investigate how irreducible, estimation and volatility uncertainty influenced the quality of action execution, here force generation, as well as the speed of action selection.

7.1.2 The probabilistic learning task

The major advantage of adopting the probabilistic learning task (PLT) (den Ouden et al., 2010) in **Chapter 6** was that it enabled direct comparisons to be made to our previous study that had used the same fundamental paradigm (de Berker et al., 2016). Importantly, by applying the same HGF model to a novel dataset acquired from participants that received a noradrenergic drug or placebo, it was possible to replicate

our earlier finding that baseline pupil diameter reflects irreducible uncertainty, and to extend previous results by establishing that pupil dilation is modulated by surprise and noradrenergic computations of volatility in a quantitatively rigorous fashion. My review of the previous pupillometric literature in **Chapter 1** highlighted the diversity of paradigmatic approaches implemented to study changes in pupil diameter under different combinations of uncertainty, surprise and volatility. A key rhetoric of this thesis is the need for researchers to adopt unified frameworks of uncertainty, surprise and volatility, such as that offered by the HGF, in order to develop comparable tasks powered to isolate the relative contributions of different neuromodulators to learning under different perceptual quantities.

In **Chapter 6**, I noted that no post-outcome pupillary response to volatility was observed under placebo. This finding contrasts with Browning et al.'s observation of post-outcome pupillary dilation in response to both surprise and volatility (Browning et al., 2015), and is likely due to the fact that volatility was not explicitly manipulated in the PLT. A fruitful extension to the current work would therefore be to investigate the impact of pharmacological manipulations of NA on pupillary responses to perceptual estimates within a modified version of the PLT that manipulated volatility over time. I would anticipate that, by modulating individuals' volatility estimates, this approach would produce larger effects on pupil diameter, thus making it possible to observe pupillary responses to volatility under placebo. I would also expect to replicate my finding that noradrenergic manipulations modulate the influence of volatility estimates on pupil diameter.

7.2 Benefits and limitations of the HGF model

The HGF is a general-purpose, heuristic Bayesian inference model. Taking inspiration from reinforcement learning schemes, the HGF seeks to overcome the computational complexity of traditional Bayesian approaches, and offers a means by which to capture differences in learning across individuals. A key advantage of the HGF is that it provides a generic hierarchical framework for individual learning that can be applied to a diverse set of behavioural paradigms (Iglesias et al., 2013; Diaconescu et al., 2014, 2017; Hauser et al., 2014; Vossel et al., 2014a, 2014b, 2015; de Berker et al., 2016; Marshall et al., 2016). Given that the brain would likely benefit from a flexible mechanism by which to modify learning and action, and therefore facilitate adaptation to novel environments characterised by unfamiliar contextual rules, this is arguably a sensible computational strategy.

7.2.1 A unified computational framework of uncertainty

In this thesis I have verified the HGF as a useful tool with which to model individual learning and behaviour under various forms of uncertainty inherent in the environment. In **Chapter 6**, I presented confirmatory evidence of the HGF's ability to capture individual learning under irreducible, estimation and volatility uncertainty. Furthermore, **Chapters 4** and **5** demonstrated that the novel instantiation of the HGF developed in **Chapter 3** could capture separable influences of uncertainty on learning and motor response modulation (Marshall et al., 2016). Importantly, the generalisable nature of the HGF means it can be used to probe learning and action across different behavioural paradigms, but within a unified computational framework of uncertainty.

7.2.2 Alternative models of learning and action

Nevertheless, a heuristic framework is not the only viable approach to modelling individual learning and action under uncertainty. An alternative strategy is to develop task-specific models.

7.2.2.1 The Forgetful Observer Model

In an ongoing collaboration with Franziska Bröker and Peter Dayan, we are seeking to test whether individual behaviour during the PSRTT might be captured more faithfully by a simpler, task-specific model. Specifically, a Forgetful Observer Model (FOM) has been built to reflect the properties of the PSRTT more closely by taking additional task-relevant information into account (Bröker, 2016). Like the HGF, the FOM features a perceptual model that captures individual learning and a response model that predicts participants' behaviour based on their perceptual beliefs. However, while the HGF is initially minimally adapted to a given behavioural task, such as the PSRTT, the FOM assumes that individuals make use of their prior knowledge of the task's structure. As such, the HGF assumes that individuals' beliefs about the transition probabilities between successive stimuli evolve independently of each other in the PSRTT, whereas the FOM implements a generative model that closely resembles the PSRTT's true generative process. On one hand, a task-specific model seems reasonable given that participants were informed before starting the task that one of four possible stimuli would be presented on each trial, meaning that the state space was constrained. Nonetheless, one could argue that the brain is not afforded the luxury of a constrained state space in real-world scenarios, instead requiring sufficient computational flexibility to account for unexpected events.

Another feature of the FOM is that it applies an exponential forgetting process to prior expectations. The forgetting rate weights past experience exponentially. As such, it offers

a heuristic means by which more recent transitions can have a greater impact on current probability estimates. The forgetting rate is conceptually related to the transition contingency learning rate, ω, in the HGF and allows the model to learn the transition probabilities within the different transition matrices implemented in the PSRTT.

Preliminary data suggest that, compared to the HGF, the FOM is superior at capturing individual learning and action in the PSRTT (Bröker, 2016). This is not necessarily surprising; one would expect a task-specific model to be more accurate, at least during initial learning. Further, a simpler model is not penalised by complexity. Whether the brain facilitates behavioural adaptation by constructing a multitude of task-specific models, and indeed whether the mechanisms underlying the FOM map onto neurophysiological processes, is another matter. In terms of physiological and behavioural relevance, both the HGF and the FOM have appealing properties. A heuristic, general-purpose model like the HGF would provide the brain with a flexible mechanism by which to adapt to any environmental context (Kumaran and Duzel, 2008; Shohamy and Wagner, 2008), but it is also possible that evolutionary experience has led the brain to fine-tune an array of environmentally valid, context-specific models of the world. In the latter case, an individual might learn the underlying rules of a novel environment by retrieving a similar pre-existing contextual model (Heckers et al., 2004) and adapting it to the current environmental parameters. An interesting avenue for future research would be to determine if and how individuals switch between different perceptual and response models when exposed to different environmental contexts (Boorman and Rushworth, 2009; Kumaran et al., 2009; Daw et al., 2011), perhaps via a hippocampal mechanism (Eichenbaum, 2000; Preston et al., 2004).

7.2.2.2 Modelling changes in volatility and metavolatility

It should also be noted that although the FOM features an heuristic forgetting rate that allows the model to learn the transition probabilities within the PSRTT's different transition matrices, it does not explicitly capture volatility uncertainty arising from contextual instability. Rather, it assumes a fixed, optimal learning process. This is in contrast to the HGF whose metavolatility parameter, ϑ, captures the rate at which volatility changes, with higher values implying a belief in a more unstable world and leading to a more variable learning rate. Given the relevance of volatility estimates to learning and action demonstrated in this thesis, it is important that future models incorporate a means by which to track volatility uncertainty since changes in this perceptual estimate indicate when and how learning rate should be adapted, with consequences for belief updating and response modulation.

A useful extension to both the HGF and the FOM would be an additional hierarchical level that allows estimates of metavolatility to change over time. Indeed, it seems reasonable that an individual's phasic volatility learning rate would vary across different environments with different degrees of volatility that change over different timescales.

7.2.2.3 Modelling change-points under environmental volatility

In both the PSRTT and PLT, participants were exposed to volatile environments characterised by discrete switches in probabilistic context. Neither the HGF nor the FOM acknowledges these punctate change-points directly. This contrasts with an alternative class of models that capture learning under volatility by modelling change-points explicitly (Adams and MacKay, 2007; Fearnhead and Liu, 2007; Nassar et al., 2010; Wilson et al., 2010, 2013). Early attempts to identify change-points using Bayesian inference relied on pre-specification of the rate at which they occur, i.e., the hazard rate. These models were limited practically by requiring the unrealistic assumption that the hazard rate was fixed and known in advance (Adams and MacKay, 2007; Fearnhead and Liu, 2007). Robert Wilson, Matthew Nassar and Joshua Gold offered an important methodological advancement by developing a hierarchical extension to earlier models that allowed the hazard rate itself to be inferred from the data, in turn facilitating identification of change-points (Wilson et al., 2010). Further developments permitted efficient Bayesian inference in volatile environments to be approximated by a computationally simple mixture of error-driven "delta" rules (Wilson et al., 2013). While the HGF does not explicitly model the change-points contained in the true generative process underlying the PSRTT and the PLT, I have demonstrated that true change-points are detected implicitly by the model as an increase in learning rate. In so doing, the HGF offers a flexible means of tracking individual learning in dynamic environments.

To summarise, in this thesis, I have validated the HGF as a useful tool with which to model individual learning and response modulation under various forms of environmental uncertainty. I have verified the capacity of the HGF's original instantiation to capture individual learning under irreducible, estimation and volatility uncertainty, and I have developed a novel instantiation of the HGF that is capable of capturing separable influences of uncertainty, prediction error and volatility on learning and action (Marshall et al., 2016). The heuristic computational framework offered by the HGF offers an important tool with which to probe the neuromodulatory mechanisms implemented by the brain to support learning and adaptive motor behaviour. The suggested extensions to the model may help to elucidate these mechanisms more precisely.

7.2.3 Relevance for machine learning

It is also worth noting that, compared to contemporary machine learning methods, humans are exceptionally good at inferring hidden probabilistic and causal relationships from limited experience. Indeed, as demonstrated in **Chapters 4**, **5** and **6**, individuals can closely track the probabilistic associations within a given environment and rapidly adapt to changes in contextual rules. Refining models of human learning and behaviour under uncertainty holds the potential to better elucidate the strategies employed by the brain to make these inferences. Since the brain offers seemingly efficient and effective solutions to the computational and implementational challenges of probabilistic inference, this approach might inspire new methodologies in machine learning and artificial intelligence. Nevertheless, the vast parameter space of heuristic models such as the HGF does present challenges. Different parameters relate to a variety of variables and testing the impact of each of them on learning, behaviour and/or pupil diameter presents a multiple comparisons problem. It is therefore important that all future computational analyses are hypothesis-driven. Indeed, this highlights an excellent case for the pre-registration of research studies, whereby researchers commit to their research predictions and methods before starting their experiments.

7.3 Combining pharmacology and neuroimaging

A number of key papers that inspired the work presented in this thesis used human functional magnetic resonance imaging (fMRI) to link the noradrenergic, cholinergic and dopaminergic systems to computations of different forms of environmental uncertainty. Indeed, blood-oxygenation-level-dependent (BOLD) activity in the noradrenergic locus coeruleus (LC) has been shown to dynamically track volatility uncertainty (Payzan-LeNestour et al., 2013), BOLD activity in the cholinergic basal forebrain reflects an individual's estimation uncertainty (specifically, their precision-weighted contingency PE), and BOLD activity in the dopaminergic midbrain correlates with individual estimates of precision-weighted sensory PE (Iglesias et al., 2013; Diaconescu et al., 2017). It is not possible to infer with certainty that activations in particular brain regions, with inhomogeneous cellular compositions, reflect the activity of specific neuromodulatory neurons. Adopting a pharmacological approach in **Chapters 4** and **6** made it possible to corroborate and extend the interpretations of these neuroimaging studies.

A recommendation for future investigations of the neuromodulatory bases of learning and/or behaviour under uncertainty would be to combine pharmacology and neuroimaging approaches (Mattay et al., 2000; Coull et al., 2001; Thiel et al., 2001, 2002;

Bentley et al., 2011; Chowdhury et al., 2013; Crockett et al., 2013). Taking advantage of the two complementary methodologies will enable researchers to pinpoint the contributions of NA, ACh and DA to different neurophysiological processes occurring in different brain regions. Future work will also be required to characterise the physiological roles of the different neuromodulators acting at different receptor sub-types and over multiple timescales. Related, continued cross-talk between human and animal research will facilitate the isolation of neuromodulatory contributions to these processes. By developing behavioural tasks and computational learning models that can be translated to different species, it will be possible to utilise a wide repertoire of methodologies, including more invasive techniques such as neuronal lesions, electrophysiological recordings and optogenetics, to characterise the specific neuromodulatory underpinnings of learning and action in uncertain environments. As recently set out by Krakauer et al., human behavioural neuroscience is well-placed to elucidate an understanding of learning and action under uncertainty through careful experimental decomposition of behaviour which, with associated pharmacological, behavioural genetics and neuroimaging approaches, can be linked to different neuromodulatory systems and brain regions. In turn, this sets the stage to causally test the precise neural implementation of behaviour using invasive interventions in animals (Krakauer et al., 2017). Nonetheless, as argued throughout this thesis, critical to this approach will be the use of unified computational models with which to interrogate the contributions of different perceptual and behavioural parameters to the neural activations revealed by neuroimaging.

7.3.1 Neural implementation of uncertainty computations

While the focus of this thesis has been to characterise the relative contributions of NA, ACh and DA to learning and response modulation under uncertainty, parallel lines of work focus on how computations of uncertainty are implemented at a neural coding level (Ma and Jazayeri, 2014; Pouget et al., 2016). Three types of neural code have been proposed (Sanger, 1996; Pouget et al., 2003; Ma et al., 2006). First, in probabilistic population coding, uncertainty is represented implicitly in neuronal population activity. For instance, increased uncertainty may correspond to a lower total spike count. Second, in sampling coding, the activity of a neuron at a given time-point is a sample from the belief distribution that is to be represented, and the probability of a variable of interest is directly mapped onto neuronal firing rate. Third, in explicit population coding, the activity of a neuron tuned to a stimulus feature is monotonically related to the probability density of that feature. Higher uncertainty is represented by a wider activation pattern across the population. The details of how the brain represents probability distributions are currently

unknown (Pouget et al., 2013), but, in addition to animal work, human neuroimaging methods such as fMRI repetition suppression hold the potential to elucidate the underlying neural mechanisms employed by the noradrenergic, cholinergic and dopaminergic systems (Barron et al., 2016). The use of fMRI will also help to isolate the precise neural networks that support uncertainty encoding (Behrens et al., 2007; Iglesias et al., 2013; Payzan-LeNestour et al., 2013; Silvetti et al., 2013; McGuire et al., 2014; Diaconescu et al., 2017).

7.4 Combining pharmacology and behavioural genetics

As I addressed in **Chapter 5**, behavioural genetics holds the potential to extend the insights offered by pharmacological and neuroimaging investigations of the neuromodulatory contributions to learning and response modulation. Importantly, a further reason to combine pharmacology and behavioural genetics is that individual behavioural responses to pharmacological manipulations can depend strongly on baseline neurotransmission (Kimberg et al., 1997; Mehta et al., 2004a; Roesch-Ely et al., 2005; Frank and O'Reilly, 2006; Cools et al., 2007b; Clatworthy et al., 2009). Since multiple genes are thought to modulate baseline neuromodulatory function, there is strong reason to predict that individual differences in noradrenergic, cholinergic and dopaminergic drug effects are, at least in part, genetic.

For instance, it has been demonstrated that the direction of cognitive effects produced by pharmacological catechol-O-methyltransferase (COMT) inhibition under tolcapone is determined by an individual's *COMT* Val^{158}Met genotype (Farrell et al., 2012). Specifically, tolcapone improves working memory performance in Val carriers but impairs performance in Met carriers. This is thought to be due to an inverted-U relationship between cortical DA neurotransmission and cognitive function (Figure 7.1). Val/Val homozygotes show higher baseline COMT activity, and thus lower baseline DA neurotransmission, than Met/Met homozygotes (Männistö and Kaakkola, 1999; Chen et al., 2004). Due to the supposed inverted-U relationship between cortical DA and cognitive function, Val/Val homozygotes show inferior baseline working memory performance compared to Met carriers. COMT inhibition under tolcapone increases cortical DA neurotransmission in all individuals but the different baseline DA levels between individuals, and the inverted-U relationship between DA and working memory, mean that the functional consequences of this shift differ between *COMT* genotypes. Val/Val individuals move closer to optimal performance, while Met individuals move past the peak, resulting in a deterioration in working memory performance.

Similarly, amphetamines, which alter catecholaminergic neurotransmission by blocking the action of transporters at dopaminergic, serotonergic and noradrenergic neurons, show variable effects in individuals with different *COMT* genotypes (Mattay et al., 2003). Moreover, pharmacological DA D2-receptor stimulation generally improves task performance in individuals with low baseline working memory span (Kimberg et al., 1997; Frank and O'Reilly, 2006), high impulsivity (Cools et al., 2007b) or low baseline DA synthesis (Cools et al., 2009), but impairs performance in those showing the opposite baseline trait. Further, the modulatory effect of atomoxetine, which upregulates catecholamines such as DA and NA, on an individual's learning rate is dependent on their learning rate at baseline (Jepma et al., 2016), potentially indicative of inter-individual genetic variations in baseline NA neurotransmission and an inverted-U relationship between NA and cognitive function.

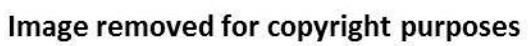

Figure 7.1 COMT genotype determines the direction of cognitive effects produced by pharmacological COMT inhibition. *An inverted-U relationship between cortical DA neurotransmission and cognitive function has been proposed. Val/Val homozygotes show higher COMT activity and lower DA neurotransmission than Met/Met homozygotes. As such, they sit further to the left on the curve at baseline, showing inferior cognitive performance. COMT inhibition under tolcapone shifts all individuals to the right because DA neurotransmission increases. The functional correlates of this shift differ between genotypes. In an n-back working memory task, Val/Val individuals move closer to optimal performance, while Met/Met individuals move to the right of the peak, resulting in declining performance. A similar effect on risk aversion is observed during a gambling task. Figure adapted from Farrell et al., 2012.*

Combining behavioural genetics and pharmacology will therefore help to improve our mechanistic understanding of the neuromodulatory contributions to learning and action in uncertain environments. The methodologies implemented in this thesis can be applied to large cohorts of healthy individuals, facilitating a refined insight into the contributions

of the different neuromodulatory systems, and the impact of different pharmacological interventions, in individuals with different genotypes. Again, the conclusions of studies adopting this strategy will only be as sophisticated as the computational models used to interrogate learning and behaviour. The approach also has clinical relevance since dopaminergic, noradrenergic and cholinergic drugs are used to treat a wide range of neurological and psychiatric disorders. This means that elucidating the relationship between baseline levels of neurotransmission and pharmacological responses will aid the development of personalised therapeutic strategies.

7.5 Functional overlaps between the noradrenergic, cholinergic and dopaminergic systems

It important to note that there are functional overlaps between the noradrenergic, cholinergic and dopaminergic systems, which limit the confidence with which the findings of this thesis can be inferred. For example, the noradrenergic LC receives inputs from several brain regions, some of them supplying dopaminergic (substantial nigra/ventral tegmental area; SN/VTA) and cholinergic (pedunculopontine tegmental nucleus; PPN) neuromodulatory influences (Samuels and Szabadi, 2008). Combining imaging of the functional activity in the SN/VTA, PPN and LC with pharmacological manipulatuons of DA and NA could offer fruitful insight by elucidating the interaction between the regions, the neuromodulatory underprinnings of these interactions, during learning and response modulation under uncertainty.

There is also considerable neurophysiological evidence that the catecholamines NA and DA have similar, partially overlapping, post-synaptic effects by boosting the efficacy of synaptic interactions between neurons, thus increasing cortical neural gain (Sutton et al., 1967; Servan-Schreiber et al., 1990; Berridge and Waterhouse, 2003; Winterer and Weinberger, 2004; Aston-Jones and Cohen, 2005a). By selectively increasing gain following unexpected outcomes, the catecholamine systems are in a position to promote belief updating in a strongly stimulus-driven manner. **Chapter 4** offered important insight into the neuromodulators' relative roles by establishing the impact of pharmacological NA, ACh *and* DA manipulations on learning and action under uncertainty. Nonetheless, the complex interactions and dependencies between the noradrenergic, cholinergic and dopaminergic systems mean that future corroborative studies using a range of drugs that target different receptor sub-types, combined with behavioural genetics and neuroimaging approaches, are required to characterise the neuromodulatory mechanisms in more detail.

7.6 Insights into neurological disorders

By characterising uncertainty computations and response modulation, the methodology reported in this thesis holds the potential to offer fresh insight into the numerous neurological and psychiatric disorders in which there is dysregulation of processes dependent on NA, ACh and DA (Iglesias et al., 2016). Further, the development of behavioural paradigms with the power to detect aberrant neuromodulatory function might offer clinically relevant diagnostic tools. As discussed previously, patients with Parkinson's disease, and therefore nigrostriatal dopaminergic depletions, show an impaired ability to make adaptive responses to unexpected sensory events that occur within a broadly predictable context and thus elicit a large sensory prediction error (Galea et al., 2012). It is possible that patients demonstrate these deficits early in their disease course (Braak et al., 2003; Anderson, 2004; Santangelo et al., 2017), meaning that tasks such as the PSRTT may offer a useful foundation for developing diagnostic behavioural markers of a dopaminergic disease state. Moreover, there has been a recent move to integrate computational neuroscience into psychiatry in an effort to better elucidate the pathophysiological mechanisms of disorders such as schizophrenia, which is also linked to (primarily cortical) dopaminergic dysfunction (Friston et al., 2014). Pinpointing the relative contributions of the neuromodulatory systems to learning and action in healthy individuals holds the potential to identify aberrant neuromodulatory processing in psychiatric disorders through assessment of learning and behaviour in tasks, such as the PSRTT and the PLT, within a unified computational framework of uncertainty, such as that offered by the HGF.

7.7 Concluding Remarks

By adopting a unified computational framework to characterise individual learning and action under irreducible, estimation and volatility uncertainty, it is possible to utilise neuropharmacology and behavioural genetics to identify the contributions of different neuromodulatory systems to perceptual belief updating and response modulation in dynamic probabilistic environments. The experiments presented in this thesis offer a foundation for future work combining pharmacology, behavioural genetics and neuroimaging to pinpoint the specific neurophysiological mechanisms by which the human brain supports learning and action in uncertain environments at cellular, network and behavioural levels.

References

Aarts E, Roelofs A, Franke B, Rijpkema M, Fernández G, Helmich RC, Cools R (2010) Striatal dopamine mediates the interface between motivational and cognitive control in humans: evidence from genetic imaging. Neuropsychopharmacology 35:1943–1951.

Abi-Dargham A, Rodenhiser J, Printz D, Zea-Ponce Y, Gil R, Kegeles LS, Weiss R, Cooper TB, Mann JJ, Van Heertum RL, others (2000) Increased baseline occupancy of D2 receptors by dopamine in schizophrenia. Proc Natl Acad Sci 97:8104–8109.

Abler B, Walter H, Erk S, Kammerer H, Spitzer M (2006) Prediction error as a linear function of reward probability is coded in human nucleus accumbens. NeuroImage 31:790–795.

Abrams P, Andersson K-E, Buccafusco JJ, Chapple C, de Groat WC, Fryer AD, Kay G, Laties A, Nathanson NM, Pasricha PJ, Wein AJ (2006) Muscarinic receptors: their distribution and function in body systems, and the implications for treating overactive bladder. Br J Pharmacol 148:565–578.

Adams RA, Huys QJM, Roiser JP (2016) Computational Psychiatry: towards a mathematically informed understanding of mental illness. J Neurol Neurosurg Psychiatry 87:53–63.

Adams RP, MacKay DJ (2007) Bayesian online changepoint detection. Tech Rep Camb. Camb Univ. Available at: https://arxiv.org/abs/0710.3742.

Akil M, Kolachana BS, Rothmond DA, Hyde TM, Weinberger DR, Kleinman JE (2003) Catechol-O-Methyltransferase Genotype and Dopamine Regulation in the Human Brain. J Neurosci 23:2008–2013.

Alexander GE, DeLong MR, Strick PL (1986) Parallel organization of functionally segregated circuits linking basal ganglia and cortex. Annu Rev Neurosci 9:357–381.

Anderson KE (2004) Behavioral disturbances in Parkinson's disease. Dialogues Clin Neurosci 6:323–332.

Arnsten AFT, Contant TA (1992) Alpha-2 adrenergic agonists decrease distractibility in aged monkeys performing the delayed response task. Psychopharmacology 108:159–169.

Astafiev SV, Snyder AZ, Shulman GL, Corbetta M (2010) Comment on "Modafinil shifts human locus coeruleus to low-tonic, high-phasic activity during functional MRI" and "Homeostatic sleep pressure and responses to sustained attention in the suprachiasmatic area." Science 328:309.

Aston-Jones G, Cohen JD (2005a) An integrative theory of locus coeruleus-norepinephrine function: adaptive gain and optimal performance. Annu Rev Neurosci 28:403–450.

Aston-Jones G, Cohen JD (2005b) Adaptive gain and the role of the locus coeruleus–norepinephrine system in optimal performance. J Comp Neurol 493:99–110.

Aston-Jones G, Rajkowski J, Kubiak P (1997) Conditioned responses of monkey locus coeruleus neurons anticipate acquisition of discriminative behavior in a vigilance task. Neuroscience 80:697–715.

Aston-Jones G, Rajkowski J, Kubiak P, Alexinsky T (1994) Locus coeruleus neurons in monkey are selectively activated by attended cues in a vigilance task. J Neurosci 14:4467–4480.

Averbeck BB, Latham PE, Pouget A (2006) Neural correlations, population coding and computation. Nat Rev Neurosci 7:358–366.

Bach DR, Dolan RJ (2012) Knowing how much you don't know: a neural organization of uncertainty estimates. Nat Rev Neurosci 13:572–586.

Bäckman L, Ginovart N, Dixon RA, Wahlin TB, Wahlin A, Halldin C, Farde L (2000) Age-related cognitive deficits mediated by changes in the striatal dopamine system. Am J Psychiatry 157:635–637.

Bäckman L, Nyberg L, Lindenberger U, Li S-C, Farde L (2006) The correlative triad among aging, dopamine, and cognition: current status and future prospects. Neurosci Biobehav Rev 30:791–807.

Balleine BW (2005) Neural bases of food-seeking: affect, arousal and reward in corticostriatolimbic circuits. Physiol Behav 86:717–730.

Bari A, Aston-Jones G (2013) Atomoxetine modulates spontaneous and sensory-evoked discharge of locus coeruleus noradrenergic neurons. Neuropharmacology 64:53–64.

Bar M (2009) Predictions: a universal principle in the operation of the human brain. Philos Trans R Soc Lond B Biol Sci 364:1181–1182.

Barron HC, Garvert MM, Behrens TEJ (2016) Repetition suppression: a means to index neural representations using BOLD? Phil Trans R Soc B 371:20150355.

Battaglia PW, Jacobs RA, Aslin RN (2003) Bayesian integration of visual and auditory signals for spatial localization. J Opt Soc Am A Opt Image Sci Vis 20:1391–1397.

Bayer HM, Glimcher PW (2005) Midbrain Dopamine Neurons Encode a Quantitative Reward Prediction Error Signal. Neuron 47:129–141.

Bays PM, Wolpert DM (2007) Computational principles of sensorimotor control that minimize uncertainty and variability. J Physiol 578:387–396.

Beal MJ (2003) Variational algorithms for approximate Bayesian inference. PhD Thesis. UCL. Available at: http://www.cse.buffalo.edu/faculty/mbeal/papers/beal03.pdf.

Beatty J (1982) Task-evoked pupillary responses, processing load, and the structure of processing resources. Psychol Bull 91:276.

Beatty WW, Monson N (1990) Problem solving in Parkinson's disease: Comparison of performance on the Wisconsin and California Card Sorting Tests. J Geriatr Psychiatry Neurol 3:163–171.

Beck JM, Ma WJ, Kiani R, Hanks T, Churchland AK, Roitman J, Shadlen MN, Latham PE, Pouget A (2008) Probabilistic population codes for Bayesian decision making. Neuron 60:1142–1152.

Behrens TEJ, Hunt LT, Woolrich MW, Rushworth MFS (2008) Associative learning of social value. Nature 456:245–249.

Behrens TEJ, Woolrich MW, Walton ME, Rushworth MFS (2007) Learning the value of information in an uncertain world. Nat Neurosci 10:1214–1221.

Beierholm U, Guitart-Masip M, Economides M, Chowdhury R, Düzel E, Dolan R, Dayan P (2013) Dopamine modulates reward-related vigor. Neuropsychopharmacology 38:1495–1503.

Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Ser B Methodol 57:289–300.

Bentley P, Driver J, Dolan RJ (2011) Cholinergic modulation of cognition: Insights from human pharmacological functional neuroimaging. Prog Neurobiol 94:360–388.

Berardelli A, Rothwell JC, Thompson PD, Hallett M (2001) Pathophysiology of bradykinesia in Parkinson's disease. Brain 124:2131–2146.

Berridge CW, Waterhouse BD (2003) The locus coeruleus–noradrenergic system: modulation of behavioral state and state-dependent cognitive processes. Brain Res Rev 42:33–84.

Bestmann S, Harrison LM, Blankenburg F, Mars RB, Haggard P, Friston KJ, Rothwell JC (2008) Influence of uncertainty and surprise on human corticospinal excitability during preparation for action. Curr Biol 18:775–780.

Bestmann S, Ruge D, Rothwell J, Galea JM (2014) The role of dopamine in motor flexibility. J Cogn Neurosci 27:365–376.

Björklund A, Dunnett SB (2007) Dopamine neuron systems in the brain: an update. Trends Neurosci 30:194–202.

Blais A-R, Weber EU (2006) A domain-specific risk-taking (DOSPERT) scale for adult populations. Judgm Decis Mak 1:33–47.

Bland AR, Schaefer A (2012) Different varieties of uncertainty in human decision-making. Front Neurosci 6:1–11.

Blasi G, Mattay VS, Bertolino A, Elvevåg B, Callicott JH, Das S, Kolachana BS, Egan MF, Goldberg TE, Weinberger DR (2005) Effect of catechol-O-methyltransferase val158met genotype on attentional control. J Neurosci 25:5038–5045.

Bond A, Lader M (1974) The use of analogue scales in rating subjective feelings. Br J Med Psychol 47:211–218.

Boorman ED, O'Doherty JP, Adolphs R, Rangel A (2013) The behavioral and neural mechanisms underlying the tracking of expertise. Neuron 80:1558–1571.

Boorman ED, Rushworth MFS (2009) Conceptual representation and the making of new decisions. Neuron 63:721–723.

Botvinick MM, Braver TS, Barch DM, Carter CS, Cohen JD (2001) Conflict monitoring and cognitive control. Psychol Rev 108:624.

Bouret S, Sara SJ (2004) Reward expectation, orientation of attention and locus coeruleus-medial frontal cortex interplay during learning. Eur J Neurosci 20:791–802.

Bouret S, Sara SJ (2005) Network reset: a simplified overarching theory of locus coeruleus noradrenaline function. Trends Neurosci 28:574–582.

Bowman EM, Brown VJ, Kertzman C, Schwarz U, Robinson DL (1993) Covert orienting of attention in macaques. I. Effects of behavioral context. J Neurophysiol 70:431–443.

Braak H, Del Tredici K, Rüb U, de Vos RAI, Jansen Steur ENH, Braak E (2003) Staging of brain pathology related to sporadic Parkinson's disease. Neurobiol Aging 24:197–211.

Bradley PB (2014) Introduction to neuropharmacology. Butterworth-Heinemann Ltd.

Bradshaw J (1967) Pupil size as a measure of arousal during information processing. Nature 216:515-516.

Bresciani J-P, Dammeier F, Ernst MO (2006) Vision and touch are automatically integrated for the perception of sequences of events. J Vis 6:554-564.

Broadbent DE, Cooper PF, FitzGerald P, Parkes KR (1982) The cognitive failures questionnaire (CFQ) and its correlates. Br J Clin Psychol 21:1–16.

Bröker F (2016) General versus special purpose inference in a complex dynamic task. Master's Thesis. UCL.

Browning M, Behrens TE, Jocham G, O'Reilly JX, Bishop SJ (2015) Anxious individuals have difficulty learning the causal statistics of aversive environments. Nat Neurosci 18:590–596.

Brown SBRE, van der Wee NJA, van Noorden MS, Giltay EJ, Nieuwenhuis S (2015) Noradrenergic and cholinergic modulation of late ERP responses to deviant stimuli. Psychophysiology 52:1620–1631.

Bucci DJ, Holland PC, Gallagher M (1998) Removal of cholinergic input to rat posterior parietal cortex disrupts incremental processing of conditioned stimuli. J Neurosci 18:8038–8046.

Bunzeck N, Düzel E (2006) Absolute coding of stimulus novelty in the human substantia nigra/VTA. Neuron 51:369–379.

Bunzeck N, Guitart-Masip M, Dolan RJ, Duzel E (2013) Pharmacological dissociation of novelty responses in the human brain. Cereb Cortex 24:1351-1360.

Bymaster FP, Katner JS, Nelson DL, Hemrick-Luecke SK, Threlkeld PG, Heiligenstein JH, Morin SM, Gehlert DR, Perry KW (2002) Atomoxetine increases extracellular levels of norepinephrine and dopamine in prefrontal cortex of rat: a potential mechanism for efficacy in attention deficit/hyperactivity disorder. Neuropsychopharmacology 27:699–711.

Caron MG (1996) Images in neuroscience. Molecular biology, II. A dopamine transporter mouse knockout. Am J Psychiatry 153:1515.

Caulfield MP (1993) Muscarinic receptors: Characterization, coupling and function. Pharmacol Ther 58:319–379.

Cavanagh JF, Sanguinetti JL, Allen JJ, Sherman SJ, Frank MJ (2014) The subthalamic nucleus contributes to post-error slowing. J Cogn Neurosci 26:2637–2644.

Chen J, Lipska BK, Halim N, Ma QD, Matsumoto M, Melhem S, Kolachana BS, Hyde TM, Herman MM, Apud J, Egan MF, Kleinman JE, Weinberger DR (2004) Functional analysis of genetic variation in catechol-O-methyltransferase (COMT): Effects on mRNA, protein, and enzyme activity in postmortem human brain. Am J Hum Genet 75:807–821.

Chen Y-CI, Choi J-K, Andersen SL, Rosen BR, Jenkins BG (2005) Mapping dopamine D2/D3 receptor function using pharmacological magnetic resonance imaging. Psychopharmacology 180:705–715.

Chiba AA, Bushnell PJ, Oshiro WM, Gallagher M (1999) Selective removal of cholinergic neurons in the basal forebrain alters cued target detection. Neuroreport 10:3119–3123.

Chowdhury R, Guitart-Masip M, Lambert C, Dayan P, Huys Q, Düzel E, Dolan RJ (2013) Dopamine restores reward prediction errors in old age. Nat Neurosci 16:648–653.

Clark CR, Geffen GM, Geffen LB (1989) Catecholamines and the covert orientation of attention in humans. Neuropsychologia 27:131–139.

Clatworthy PL, Lewis SJG, Brichard L, Hong YT, Izquierdo D, Clark L, Cools R, Aigbirhio FI, Baron J-C, Fryer TD, Robbins TW (2009) Dopamine release in dissociable striatal subregions predicts the different effects of oral methylphenidate on reversal learning and spatial working memory. J Neurosci 29:4690–4696.

Cohen MX, Krohn-Grimberghe A, Elger CE, Weber B (2007) Dopamine gene predicts the brain's response to dopaminergic drug. Eur J Neurosci 26:3652–3660.

Conant RC, Ashby WR (1970) Every good regulator of a system must be a model of that system. Int J Syst Sci 1:89–97.

Cools AR, van den Bercken JH, Horstink MW, Van Spaendonck KP, Berger HJ (1984) Cognitive and motor shifting aptitude disorder in Parkinson's disease. J Neurol Neurosurg Psychiatry 47:443–453.

Cools R, Barker RA, Sahakian BJ, Robbins TW (2001a) Mechanisms of cognitive set flexibility in Parkinson's disease. Brain 124:2503–2512.

Cools R, Barker RA, Sahakian BJ, Robbins TW (2001b) Enhanced or Impaired Cognitive Function in Parkinson's Disease as a Function of Dopaminergic Medication and Task Demands. Cereb Cortex 11:1136–1143.

Cools R, Frank MJ, Gibbs SE, Miyakawa A, Jagust W, D'Esposito M (2009) Striatal dopamine predicts outcome-specific reversal learning and its sensitivity to dopaminergic drug administration. J Neurosci 29:1538–1543.

Cools R, Lewis SJG, Clark L, Barker RA, Robbins TW (2007a) L-DOPA disrupts activity in the nucleus accumbens during reversal learning in Parkinson's disease. Neuropsychopharmacol 32:180–189.

Cools R, Robbins TW (2004) Chemistry of the adaptive mind. Philos Transact A Math Phys Eng Sci 362:2871–2888.

Cools R, Sheridan M, Jacobs E, D'Esposito M (2007b) Impulsive personality predicts dopamine-dependent changes in frontostriatal activity during component processes of working memory. J Neurosci 27:5506–5514.

Cooper JC, Kreps TA, Wiebe T, Pirkl T, Knutson B (2010) When giving is good: ventromedial prefrontal cortex activation for others' intentions. Neuron 67:511–521.

Cooper JR, Bloom FE, Roth RH (2003) The biochemical basis of neuropharmacology. Oxford University Press, USA.

Corbetta M, Patel G, Shulman GL (2008) The reorienting system of the human brain: from environment to theory of mind. Neuron 58:306–324.

Costa A, Riedel M, Müller U, Möller H-J, Ettinger U (2011) Relationship between SLC6A3 genotype and striatal dopamine transporter availability: A meta-analysis of human single photon emission computed tomography studies. Synapse 65:998–1005.

Coull JT (2001) Modulation of attention by noradrenergic a 2-agents varies according to arousal level. Drug News Perspect 14:5-11.

Coull JT, Nobre AC, Frith CD (2001) The noradrenergic α2 agonist clonidine modulates behavioural and neuroanatomical correlates of human attentional orienting and alerting. Cereb Cortex 11:73–84.

Coull JT, Sahakian BJ, Middleton HC, Young AH, Park SB, McShane RH, Cowen PJ, Robbins TW (1995) Differential effects of clonidine, haloperidol, diazepam and tryptophan depletion on focused attention and attentional search. Psychopharmacology 121:222–230.

Critchley HD (2005) Neural mechanisms of autonomic, affective, and cognitive integration. J Comp Neurol 493:154–166.

Crockett MJ, Apergis-Schoute A, Herrmann B, Lieberman MD, Lieberman M, Müller U, Robbins TW, Clark L (2013) Serotonin modulates striatal responses to fairness and retaliation in humans. J Neurosci 33:3505–3513.

Crockett MJ, Siegel JZ, Kurth-Nelson Z, Ousdal OT, Story G, Frieband C, Grosse-Rueskamp JM, Dayan P, Dolan RJ (2015) Dissociable effects of serotonin and dopamine on the valuation of harm in moral decision making. Curr Biol 25:1852–1859.

D'Ardenne K, McClure SM, Nystrom LE, Cohen JD (2008) BOLD responses reflecting dopaminergic signals in the human ventral tegmental area. Science 319:1264–1267.

Daunizeau J, Adam V, Rigoux L (2014) VBA: a probabilistic treatment of nonlinear models for neurobiological and behavioural data. PLOS Comput Biol 10:e1003441.

Daunizeau J, den Ouden HEM, Pessiglione M, Kiebel SJ, Stephan KE, Friston KJ (2010a) Observing the observer (I): Meta-Bayesian models of learning and decision-making. PLOS ONE 5:e15554.

Daunizeau J, Ouden HEM den, Pessiglione M, Kiebel SJ, Friston KJ, Stephan KE (2010b) Observing the observer (II): Deciding when to decide. PLOS ONE 5:e15555.

Daw ND, Doya K (2006) The computational neurobiology of learning and reward. Curr Opin Neurobiol 16:199–204.

Daw ND, Gershman SJ, Seymour B, Dayan P, Dolan RJ (2011) Model-based influences on humans' choices and striatal prediction errors. Neuron 69:1204–1215.

Daw ND, O'Doherty JP (2013) Neuroeconomics: decision making and the brain. Academic Press.

Dayan P (2012) Twenty-five lessons from computational neuromodulation. Neuron 76:240–256.

Dayan P, Huys QJ (2009) Serotonin in affective control. Neuroscience 32:95.

Dayan P, Niv Y (2008) Reinforcement learning: the good, the bad and the ugly. Curr Opin Neurobiol 18:185–196.

Dayan P, Yu A (2006a) Norepinephrine and neural interrupts. Adv Neural Inf Process Syst 18:243.

Dayan P, Yu AJ (2006b) Phasic norepinephrine: a neural interrupt signal for unexpected events. Netw Comput Neural Syst 17:335–350.

de Berker AO, Rutledge RB, Mathys C, Marshall L, Cross GF, Dolan RJ, Bestmann S (2016) Computations of uncertainty mediate acute stress responses in humans. Nat Commun 7:10996.

Deco G, Thiele A (2011) Cholinergic control of cortical network interactions enables feedback-mediated attentional modulation. Eur J Neurosci 34:146–157.

de Gee JW, Knapen T, Donner TH (2014) Decision-related pupil dilation reflects upcoming choice and individual bias. Proc Natl Acad Sci 111:E618–E625.

de Martino B, Strange BA, Dolan RJ (2007) Noradrenergic neuromodulation of human attention for emotional and neutral stimuli. Psychopharmacology 197:127–136.

Deneve S (2007) Bayesian spiking neurons II: Learning. Neural Comput 20:118–145.

den Ouden HE, Daunizeau J, Roiser J, Friston KJ, Stephan KE (2010) Striatal prediction error modulates cortical coupling. J Neurosci 30:3210–3219.

den Ouden HE, Daw ND, Fernandez G, Elshout JA, Rijpkema M, Hoogman M, Franke B, Cools R (2013) Dissociable effects of dopamine and serotonin on reversal learning. Neuron 80:1090–1100.

den Ouden HEM, Friston KJ, Daw ND, McIntosh AR, Stephan KE (2009) A Dual Role for Prediction Error in Associative Learning. Cereb Cortex 19:1175–1185.

den Ouden HEM, Kok P, de Lange FP (2012) How prediction errors shape perception, attention, and motivation. Front Psychol 3:548.

den Ouden HEM, Swart JC, Schmidt K, Fekkes D, Geurts DEM, Cools R (2015) Acute serotonin depletion releases motivated inhibition of response vigour. Psychopharmacology 232:1303–1312.

de Rover M, Brown SBRE, Boot N, Hajcak G, van Noorden MS, van der Wee NJA, Nieuwenhuis S (2012) Beta receptor-mediated modulation of the late positive potential in humans. Psychopharmacology 219:971–979.

Devauges V, Sara SJ (1990) Activation of the noradrenergic system facilitates an attentional shift in the rat. Behav Brain Res 39:19–28.

Devoto P, Flore G (2006) On the origin of cortical dopamine: is it a co-transmitter in noradrenergic neurons? Curr Neuropharmacol 4:115–125.

Diaconescu AO, Mathys C, Weber LAE, Daunizeau J, Kasper L, Lomakina EI, Fehr E, Stephan KE (2014) Inferring on the Intentions of Others by Hierarchical Bayesian Learning. PLOS Comput Biol 10:e1003810.

Diaconescu AO, Mathys C, Weber LAE, Kasper L, Mauer J, Stephan KE (2017) Hierarchical prediction errors in midbrain and septum during social learning. Soc Cogn Affect Neurosci. Epub ahead of print.

Diamond A (2007) Consequences of variations in genes that affect dopamine in prefrontal cortex. Cereb Cortex 17:i161–i170.

Diaz-Asper CM, Goldberg TE, Kolachana BS, Straub RE, Egan MF, Weinberger DR (2008) Genetic variation in catechol-O-methyltransferase: effects on working memory in schizophrenic patients, their siblings, and healthy controls. Biol Psychiatry 63:72–79.

Diederen KMJ, Ziauddeen H, Vestergaard MD, Spencer T, Schultz W, Fletcher PC (2017) Dopamine Modulates Adaptive Prediction Error Coding in the Human Midbrain and Striatum. J Neurosci 37:1708–1720.

Dolan RJ (2007) The human amygdala and orbital prefrontal cortex in behavioural regulation. Philos Trans R Soc Lond B Biol Sci 362:787–799.

Doll BB, Bath KG, Daw ND, Frank MJ (2016) Variability in dopamine genes dissociates model-based and model-free reinforcement learning. J Neurosci 36:1211–1222.

Dostert P, Benedetti MS, Poggesi I (1997) Review of the pharmacokinetics and metabolism of reboxetine, a selective noradrenaline reuptake inhibitor. Eur Neuropsychopharmacol 7:S23–S35.

Downing CJ (1988) Expectancy and visual-spatial attention: Effects on perceptual quality. J Exp Psychol Hum Percept Perform 14:188–202.

Doya K (2002) Metalearning and neuromodulation. Neural Netw 15:495–506.

Doya K (2008) Modulators of decision making. Nat Neurosci 11:410–416.

Dreher J-C, Kohn P, Kolachana B, Weinberger DR, Berman KF (2009) Variation in dopamine genes influences responsivity of the human reward system. Proc Natl Acad Sci 106:617–622.

Dreher J-C, Meyer-Lindenberg A, Kohn P, Berman KF (2008) Age-related changes in midbrain dopaminergic regulation of the human reward system. Proc Natl Acad Sci U S A 105:15106–15111.

Dumontheil I, Roggeman C, Ziermans T, Peyrard-Janvid M, Matsson H, Kere J, Klingberg T (2011) Influence of the COMT genotype on working memory and brain activity changes during development. Biol Psychiatry 70:222–229.

Ebitz RB, Platt ML (2015) Neuronal activity in primate dorsal anterior cingulate cortex signals task conflict and predicts adjustments in pupil-linked arousal. Neuron 85:628–640.

Egan MF, Goldberg TE, Kolachana BS, Callicott JH, Mazzanti CM, Straub RE, Goldman D, Weinberger DR (2001) Effect of COMT Val108/158 Met genotype on frontal lobe function and risk for schizophrenia. Proc Natl Acad Sci 98:6917–6922.

Eichenbaum H (2000) A cortical–hippocampal system for declarative memory. Nat Rev Neurosci 1:41–50.

Einhauser W, Koch C, Carter O (2010) Pupil dilation betrays the timing of decisions. Front Hum Neurosci 4:18.

Einhäuser W, Stout J, Koch C, Carter O (2008) Pupil dilation reflects perceptual selection and predicts subsequent stability in perceptual rivalry. Proc Natl Acad Sci 105:1704–1709.

Eisenhofer G, Kopin IJ, Goldstein DS (2004) Catecholamine metabolism: a contemporary view with implications for physiology and medicine. Pharmacol Rev 56:331–349.

Eldar E, Cohen JD, Niv Y (2013) The effects of neural gain on attention and learning. Nat Neurosci 16:1146–1153.

Ellis BW, Johns MW, Lancaster R, Raptopoulos P, Angelopoulos N, Priest RG (1980) The St. Mary's Hospital sleep questionnaire: a study of reliability. Sleep 4:93–97.

Eppinger B, Schuck NW, Nystrom LE, Cohen JD (2013) Reduced striatal responses to reward prediction errors in older compared with younger adults. J Neurosci 33:9905–9912.

Ernst MO, Banks MS (2002) Humans integrate visual and haptic information in a statistically optimal fashion. Nature 415:429–433.

Farrell SM, Tunbridge EM, Braeutigam S, Harrison PJ (2012) COMT Val158Met genotype determines the direction of cognitive effects produced by catechol-O-methyltransferase inhibition. Biol Psychiatry 71:538–544.

Fearnhead P, Liu Z (2007) On-line inference for multiple changepoint problems. J R Stat Soc Ser B Stat Methodol 69:589–605.

Fetsch CR, Kiani R, Shadlen MN (2014) Predicting the accuracy of a decision: A neural mechanism of confidence. In: Cold Spring Harbor symposia on quantitative biology, pp 185–197.

Fiedler S, Glöckner A (2012) The dynamics of decision making in risky choice: an eye-tracking analysis. Front Psychol 3:335.

Fiser J, Berkes P, Orbán G, Lengyel M (2010) Statistically optimal perception and learning: from behavior to neural representations. Trends Cogn Sci 14:119–130.

Floderus Y, Ross SB, Wetterberg L (1981) Erythrocyte catechol-O-methyltransferase activity in a Swedish population. Clin Genet 19:389–392.

Foltynie T, Goldberg TE, Lewis SG, Blackwell AD, Kolachana BS, Weinberger DR, Robbins TW, Barker RA (2004) Planning ability in Parkinson's disease is influenced by the COMT Val158Met polymorphism. Mov Disord 19:885–891.

Forbes EE, Brown SM, Kimak M, Ferrell RE, Manuck SB, Hariri AR (2009) Genetic variation in components of dopamine neurotransmission impacts ventral striatal reactivity associated with impulsivity. Mol Psychiatry 14:60–70.

Fotiou DF, Stergiou V, Tsiptsios D, Lithari C, Nakou M, Karlovasitou A (2009) Cholinergic deficiency in Alzheimer's and Parkinson's disease: Evaluation with pupillometry. Int J Psychophysiol 73:143–149.

Franke B, Vasquez AA, Johansson S, Hoogman M, Romanos J, Boreatti-Hümmer A, Heine M, Jacob CP, Lesch K-P, Casas M, others (2010) Multicenter analysis of the SLC6A3/DAT1 VNTR haplotype in persistent ADHD suggests differential involvement of the gene in childhood and persistent ADHD. Neuropsychopharmacology 35:656–664.

Frank MJ (2005) Dynamic dopamine modulation in the basal ganglia: a neurocomputational account of cognitive deficits in medicated and nonmedicated Parkinsonism. J Cogn Neurosci 17:51–72.

Frank MJ (2008) Schizophrenia: a computational reinforcement learning perspective. Schizophr Bull 34:1008–1011.

Frank MJ, Doll BB, Oas-Terpstra J, Moreno F (2009) Prefrontal and striatal dopaminergic genes predict individual differences in exploration and exploitation. Nat Neurosci 12:1062–1068.

Frank MJ, Fossella JA (2011) Neurogenetics and pharmacology of learning, motivation, and cognition. Neuropsychopharmacology 36:133–152.

Frank MJ, Moustafa AA, Haughey HM, Curran T, Hutchison KE (2007) Genetic triple dissociation reveals multiple roles for dopamine in reinforcement learning. Proc Natl Acad Sci 104:16311–16316.

Frank MJ, O'Reilly RC (2006) A mechanistic account of striatal dopamine function in human cognition: psychopharmacological studies with cabergoline and haloperidol. Behav Neurosci 120:497.

Frias C, Annerbrink K, Westberg L, Eriksson E, Adolfsson R, Nilsson L (2005) Catechol O-methyltransferase Val158Met polymorphism is associated with cognitive performance in nondemented adults. J Cogn Neurosci 17:1018–1025.

Friedman NP, Miyake A, Young SE, DeFries JC, Corley RP, Hewitt JK (2008) Individual differences in executive functions are almost entirely genetic in origin. J Exp Psychol Gen 137:201.

Friston K (2005) A theory of cortical responses. Philos Trans R Soc B Biol Sci 360:815–836.

Friston KJ, Shiner T, FitzGerald T, Galea JM, Adams R, Brown H, Dolan RJ, Moran R, Stephan KE, Bestmann S (2012) Dopamine, affordance and active inference. PLOS Comput Biol 8:e1002327.

Friston KJ, Stephan KE (2007) Free-energy and the brain. Synthese 159:417–458.

Friston KJ, Stephan KE, Montague R, Dolan RJ (2014) Computational psychiatry: the brain as a phantastic organ. Lancet Psychiatry 1:148–158.

Friston K, Kilner J, Harrison L (2006) A free energy principle for the brain. J Physiol-Paris 100:70–87.

Fuke S, Suo S, Takahashi N, Koike H, Sasagawa N, Ishiura S (2001) The VNTR polymorphism of the human dopamine transporter (DAT1) gene affects gene expression. Pharmacogenomics J 1:152–156.

Galea JM, Bestmann S, Beigi M, Jahanshahi M, Rothwell JC (2012) Action reprogramming in Parkinson's disease: response to prediction error is modulated by levels of dopamine. J Neurosci 32:542–550.

Galea JM, Ruge D, Buijink A, Bestmann S, Rothwell JC (2013) Punishment-induced behavioral and neurophysiological variability reveals dopamine-dependent selection of kinematic movement parameters. J Neurosci 33:3981–3988.

Garcia-Garcia M, Barcelo F, Clemente IC, Escera C (2010) The role of the dopamine transporter DAT1 genotype on the neural correlates of cognitive flexibility. Eur J Neurosci 31:754–760.

Garrison E, Marth G (2012) Haplotype-based variant detection from short-read sequencing. Available at: http://arxiv.org/abs/1207.3907.

Garris PA, Budygin EA, Phillips PEM, Venton BJ, Robinson DL, Bergstrom BP, Rebec GV, Wightman RM (2003) A role for presynaptic mechanisms in the actions of nomifensine and haloperidol. Neuroscience 118:819–829.

Gehring WJ, Fencsik DE (2001) Functions of the medial frontal cortex in the processing of conflict and errors. J Neurosci 21:9430–9437.

Geisler WS, Diehl RL (2002) Bayesian natural selection and the evolution of perceptual systems. Philos Trans R Soc Lond B Biol Sci 357:419–448.

German DC, Walker BS, Manaye K, Smith WK, Woodward DJ, North AJ (1988) The human locus coeruleus: computer reconstruction of cellular distribution. J Neurosci 8:1776–1788.

Gershman SJ, Niv Y (2010) Learning latent structure: carving nature at its joints. Curr Opin Neurobiol 20:251–256.

Gil Z, Connors BW, Amitai Y (1997) Differential regulation of neocortical synapses by neuromodulators and activity. Neuron 19:679–686.

Gilzenrat MS, Nieuwenhuis S, Jepma M, Cohen JD (2010) Pupil diameter tracks changes in control state predicted by the adaptive gain theory of locus coeruleus function. Cogn Affect Behav Neurosci 10:252–269.

Gingrich JA, Caron MG (1993) Recent advances in the molecular biology of dopamine receptors. Annu Rev Neurosci 16:299–321.

Gizer IR, Ficks C, Waldman ID (2009) Candidate gene studies of ADHD: a meta-analytic review. Hum Genet 126:51–90.

Gluck MA, Shohamy D, Myers C (2002) How do people solve the “weather prediction” task?: Individual variability in strategies for probabilistic category learning. Learn Mem 9:408–418.

Gogos JA, Morgan M, Luine V, Santha M, Ogawa S, Pfaff D, Karayiorgou M (1998) Catechol-O-methyltransferase-deficient mice exhibit sexually dimorphic changes in catecholamine levels and behavior. Proc Natl Acad Sci 95:9991–9996.

Goldberg T, Egan M, Gscheidle T, et al (2003) Executive subprocesses in working memory: Relationship to catechol-o-methyltransferase Val158Met genotype and schizophrenia. Arch Gen Psychiatry 60:889–896.

Goldman-Rakic PS (1998) The cortical dopamine system: role in memory and cognition. Adv Pharmacol 42:707–711.

Gorman JM, Hirschfeld RM, Ninan PT (2001) New developments in the neurobiological basis of anxiety disorders. Psychopharmacol Bull 36 Suppl 2:49–67.

Goto Y, Grace AA (2005) Dopaminergic modulation of limbic and cortical drive of nucleus accumbens in goal-directed behavior. Nat Neurosci 8:805–812.

Grace AA (1995) The tonic/phasic model of dopamine system regulation: its relevance for understanding how stimulant abuse can alter basal ganglia function. Drug Alcohol Depend 37:111–129.

Grace AA (2000) Gating of information flow within the limbic system and the pathophysiology of schizophrenia. Brain Res Rev 31:330–341.

Graham DG (1979) On the origin and significance of neuromelanin. Arch Pathol Lab Med 103:359–362.

Granon S, Passetti F, Thomas KL, Dalley JW, Everitt BJ, Robbins TW (2000) Enhanced and Impaired attentional performance after infusion of D1 dopaminergic receptor agents into rat prefrontal cortex. J Neurosci 20:1208–1215.

Green AE, Munafò MR, DeYoung CG, Fossella JA, Fan J, Gray JR (2008) Using genetic data in cognitive neuroscience: from growing pains to genuine insights. Nat Rev Neurosci 9:710–720.

Green DM (1993) A maximum-likelihood method for estimating thresholds in a yes–no task. J Acoust Soc Am 93:2096–2105.

Gregory RL (1970) The Intelligent Eye. Littlehampton Book Services Ltd.

Gregory RL (1997) Knowledge in perception and illusion. Philos Trans R Soc Lond B Biol Sci 352:1121–1127.

Guitart-Masip M, Duzel E, Dolan R, Dayan P (2014) Action versus valence in decision making. Trends Cogn Sci 18:194–202.

Gu Q (2002) Neuromodulatory transmitter systems in the cortex and their role in cortical plasticity. Neuroscience 111:815–835.

Haber SN, Fudge JL, McFarland NR (2000) Striatonigrostriatal pathways in primates form an ascending spiral from the shell to the dorsolateral striatum. J Neurosci 20:2369–2382.

Hall H, Sedvall G, Magnusson O, Kopp J, Halldin C, Farde L (1994) Distribution of D1- and D2-dopamine receptors, and dopamine and its metabolites in the human brain. Neuropsychopharmacology 11:245–256.

Hampton AN, Bossaerts P, O'doherty JP (2006) The role of the ventromedial prefrontal cortex in abstract state-based inference during decision making in humans. J Neurosci 26:8360–8367.

Hare TA, O'Doherty J, Camerer CF, Schultz W, Rangel A (2008) Dissociating the role of the orbitofrontal cortex and the striatum in the computation of goal values and prediction errors. J Neurosci 28:5623–5630.

Harley CW (1987) A role for norepinephrine in arousal, emotion and learning?: limbic modulation by norepinephrine and the Kety hypothesis. Prog Neuropsychopharmacol Biol Psychiatry 11:419–458.

Harrison LM, Duggins A, Friston KJ (2006) Encoding uncertainty in the hippocampus. Neural Netw 19:535–546.

Hart AS, Rutledge RB, Glimcher PW, Phillips PEM (2014) Phasic dopamine release in the rat nucleus accumbens symmetrically encodes a reward prediction error term. J Neurosci 34:698–704.

Hasselmo ME (1995) Neuromodulation and cortical function: modeling the physiological basis of behavior. Behav Brain Res 67:1–27.

Hasselmo ME, McGaughy J (2004) High acetylcholine levels set circuit dynamics for attention and encoding and low acetylcholine levels set dynamics for consolidation. Prog Brain Res 145:207-231.

Hasselmo ME, Wyble BP, Wallenstein GV (1996) Encoding and retrieval of episodic memories: role of cholinergic and GABAergic modulation in the hippocampus. Hippocampus 6:693–708.

Hauser TU, Iannaccone R, Ball J, Mathys C, Brandeis D, Walitza S, Brem S (2014) Role of the medial prefrontal cortex in impaired decision making in juvenile attention-deficit/hyperactivity disorder. JAMA Psychiatry 71:1165–1173.

Heckers S, Zalesak M, Weiss AP, Ditman T, Titone D (2004) Hippocampal activation during transitive inference in humans. Hippocampus 14:153–162.

Heinz A, Goldman D, Jones DW, Palmour R, Hommer D, Gorey JG, Lee KS, Linnoila M, Weinberger DR (2000) Genotype influences in vivo dopamine transporter availability in human striatum. Neuropsychopharmacology 22:133–139.

Heinz A, Saunders RC, Kolachana BS, Bertolino A, Jones DW, Gorey JG, Bachevalier J, Lee KS, Knable MB, Saunders RC, others (1999) Disinhibition of subcortical dopaminergic neurotransmission in rhesus monkeys with neonatal mesial temporal lesions. Synapse 39:71–79.

Hess EH, Polt JM (1964) Pupil size in relation to mental activity during simple problem-solving. Science 143:1190–1192.

Hick WE (1952) On the rate of gain of information. Q J Exp Psychol 4:11–26.

Hikosaka O, Isoda M (2010) Switching from automatic to controlled behavior: cortico-basal ganglia mechanisms. Trends Cogn Sci 14:154–161.

Himmelheber AM, Sarter M, Bruno JP (2000) Increases in cortical acetylcholine release during sustained attention performance in rats. Brain Res Cogn Brain Res 9:313–325.

Hollerman JR, Schultz W (1998) Dopamine neurons report an error in the temporal prediction of reward during learning. Nat Neurosci 1:304–309.

Houk JC, Adams JL, Barto AG (1995) A model of how the basal ganglia generate and use neural signals that predict reinforcement. In Houk JC, Davis JL, Beiser DG. Models of information processing in the basal ganglia. MIT Press.

Huang S-Y, Lu R-B, Ma K-H, Shy M-J, Lin W-W (2008) Norepinephrine transporter polymorphisms T-182C and G1287A are not associated with alcohol dependence and its clinical subgroups. Drug Alcohol Depend 92:20–26.

Hupé J-M, Lamirel C, Lorenceau J (2009) Pupil dynamics during bistable motion perception. J Vis 9:10.

Hyman R (1953) Stimulus information as a determinant of reaction time. J Exp Psychol 45:188.

Iglesias S, Mathys C, Brodersen KH, Kasper L, Piccirelli M, den Ouden HEM, Stephan KE (2013) Hierarchical prediction errors in midbrain and basal forebrain during sensory learning. Neuron 80:519–530.

Iglesias S, Tomiello S, Schneebeli M, Stephan KE (2016) Models of neuromodulation for computational psychiatry. WIREs Cogn Sci.

Invernizzi RW, Garattini S (2004) Role of presynaptic α2-adrenoceptors in antidepressant action: recent findings from microdialysis studies. Prog Neuropsychopharmacol Biol Psychiatry 28:819–827.

Isoda M, Hikosaka O (2011) Cortico-basal ganglia mechanisms for overcoming innate, habitual and motivational behaviors. Eur J Neurosci 33:2058–2069.

Itami S, Uno H (2002) Orbitofrontal cortex dysfunction in attention-deficit hyperactivity disorder revealed by reversal and extinction tasks. Neuroreport 13:2453–2457.

Jacobsen LK, Staley JK, Zoghbi SS, Seibyl JP, Kosten TR, Innis RB, Gelernter J (2000) Prediction of dopamine transporter binding availability by genotype: a preliminary report. Am J Psychiatry 157:1700-1703.

Jeannotte AM, McCarthy JG, Sidhu A (2009) Desipramine induced changes in the norepinephrine transporter, alpha- and gamma-synuclein in the hippocampus, amygdala and striatum. Neurosci Lett 467:86–89.

Jepma M, Murphy PR, Nassar MR, Rangel-Gomez M, Meeter M, Nieuwenhuis S (2016) Catecholaminergic regulation of learning rate in a dynamic environment. PLOS Comput Biol 12:e1005171.

Jepma M, Nieuwenhuis S (2011) Pupil diameter predicts changes in the exploration–exploitation trade-off: evidence for the adaptive gain theory. J Cogn Neurosci 23:1587–1596.

Jepma M, Te Beek ET, Wagenmakers E-J, Van Gerven JMA, Nieuwenhuis S (2010) The role of the noradrenergic system in the exploration–exploitation trade-off: a psychopharmacological study. Front Hum Neurosci 4:170.

Jones BE (2005) From waking to sleeping: neuronal and chemical substrates. Trends Pharmacol Sci 26:578–586.

Jones DNC, Higgins GA (1995) Effect of scopolamine on visual attention in rats. Psychopharmacology 120:142–149.

Jones LS, Gauger LL, Davis JN (1985) Anatomy of brain alpha 1-adrenergic receptors: in vitro autoradiography with [125I]-heat. J Comp Neurol 231:190–208.

Joshi S, Li Y, Kalwani RM, Gold JI (2016) Relationships between pupil diameter and neuronal activity in the locus coeruleus, colliculi, and cingulate cortex. Neuron 89:221–234.

Joshua M, Adler A, Mitelman R, Vaadia E, Bergman H (2008) Midbrain dopaminergic neurons and striatal cholinergic interneurons encode the difference between reward and aversive events at different epochs of probabilistic classical conditioning trials. J Neurosci 28:11673–11684.

Kaasinen V, Vilkman H, Hietala J, Någren K, Helenius H, Olsson H, Farde L, Rinne J (2000) Age-related dopamine D2/D3 receptor loss in extrastriatal regions of the human brain. Neurobiol Aging 21:683–688.

Kahneman D, Beatty J (1966) Pupil diameter and load on memory. Science 154:1583–1585.

Kahneman D, Beatty J, Pollack I (1967) Perceptual deficit during a mental task. Science 157:218–219.

Kalman RE (1960) A new approach to linear filtering and prediction problems. J Basic Eng 82:35–45.

Kehagia AA, Murray GK, Robbins TW (2010) Learning and cognitive flexibility: frontostriatal function and monoaminergic modulation. Curr Opin Neurobiol 20:199–204.

Keren NI, Lozar CT, Harris KC, Morgan PS, Eckert MA (2009) In vivo mapping of the human locus coeruleus. NeuroImage 47:1261–1267.

Killcross S, Coutureau E (2003) Coordination of actions and habits in the medial prefrontal cortex of rats. Cereb Cortex 13:400–408.

Kimberg DY, D'Esposito M, Farah MJ (1997) Effects of bromocriptine on human subjects depend on working memory capacity. Neuroreport 8:3581–3585.

Kimura F, Fukuda M, Tsumoto T (1999) Acetylcholine suppresses the spread of excitation in the visual cortex revealed by optical recording: possible differential effect depending on the source of input. Eur J Neurosci 11:3597–3609.

Klein-Flügge MC, Hunt LT, Bach DR, Dolan RJ, Behrens TE (2011) Dissociable reward and timing signals in human midbrain and ventral striatum. Neuron 72:654–664.

Knill DC, Pouget A (2004) The Bayesian brain: the role of uncertainty in neural coding and computation. Trends Neurosci 27:712–719.

Kobayashi M, Imamura K, Sugai T, Onoda N, Yamamoto M, Komai S, Watanabe Y (2000) Selective suppression of horizontal propagation in rat visual cortex by norepinephrine. Eur J Neurosci 12:264–272.

Koda K, Ago Y, Cong Y, Kita Y, Takuma K, Matsuda T (2010) Effects of acute and chronic administration of atomoxetine and methylphenidate on extracellular levels of noradrenaline, dopamine and serotonin in the prefrontal cortex and striatum of mice. J Neurochem 114:259–270.

Korchounov A, Ziemann U (2011) Neuromodulatory neurotransmitters influence LTP-like plasticity in human cortex: a pharmaco-TMS study. Neuropsychopharmacology 36:1894–1902.

Körding K (2007) Decision theory: what "should" the nervous system do? Science 318:606–610.

Körding KP, Wolpert DM (2004) Bayesian integration in sensorimotor learning. Nature 427:244–247.

Korn CW, Staib M, Tzovara A, Castegnetti G, Bach DR (2016) A pupil size response model to assess fear learning. Psychophysiology 54:330-343.

Krakauer JW, Ghazanfar AA, Gomez-Marin A, MacIver MA, Poeppel D (2017) Neuroscience needs behavior: Correcting a reductionist bias. Neuron 93:480–490.

Kumaran D, Duzel E (2008) The Hippocampus and Dopaminergic Midbrain: Old Couple, New Insights. Neuron 60:197–200.

Kumaran D, Summerfield JJ, Hassabis D, Maguire EA (2009) Tracking the Emergence of Conceptual Knowledge during Human Decision Making. Neuron 63:889–901.

Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. Nat Methods 9:357–359.

Lapiz MDS, Morilak DA (2006) Noradrenergic modulation of cognitive function in rat medial prefrontal cortex as measured by attentional set shifting capability. Neuroscience 137:1039–1049.

Laplace P (1774) Mémoire sur la probabilité des causes par les évènemens. Mém Acad Roy Sci 6:621–656.

Laplace P (1812) Théorie analytique des probabilités. Paris Courcier Impr 7.

Lewis DA, Melchitzky DS, Sesack SR, Whitehead RE, Auh S, Sampson A (2001) Dopamine transporter immunoreactivity in monkey cerebral cortex: regional, laminar, and ultrastructural localization. J Comp Neurol 432:119–136.

Li S-C, Lindenberger U, Bäckman L (2010) Dopaminergic modulation of cognition across the life span. Neurosci Biobehav Rev 34:625–630.

Little JT, Johnson DN, Minichiello M, Weingartner H, Sunderland T (1998) Combined nicotinic and muscarinic blockade in elderly normal volunteers: Cognitive, behavioral, and physiologic responses. Neuropsychopharmacology 19:60–69.

Malhotra AK, Kestler LJ, Mazzanti C, Bates JA, Goldberg T, Goldman D (2002) A functional polymorphism in the COMT gene and performance on a test of prefrontal cognition. Am J Psychiatry 159:652–654.

Männistö PT, Kaakkola S (1999) Catechol-O-methyltransferase (COMT): biochemistry, molecular biology, pharmacology, and clinical efficacy of the new selective COMT inhibitors. Pharmacol Rev 51:593–628.

Marder E (2012) Neuromodulation of neuronal circuits: Back to the future. Neuron 76:1–11.

Marshall L, Mathys C, Ruge D, Berker AO de, Dayan P, Stephan KE, Bestmann S (2016) Pharmacological fingerprints of contextual uncertainty. PLOS Biol 14:e1002575.

Mathys C, Daunizeau J, Friston KJ, Stephan KE (2011) A Bayesian foundation for individual learning under uncertainty. Front Hum Neurosci 5:39.

Mathys CD, Lomakina EI, Daunizeau J, Iglesias S, Brodersen KH, Friston KJ, Stephan KE (2014) Uncertainty in perception and the Hierarchical Gaussian Filter. Front Hum Neurosci 8:1–24.

Matsumoto K, Tanaka K (2004) The role of the medial prefrontal cortex in achieving goals. Curr Opin Neurobiol 14:178–185.

Matsumoto M, Hikosaka O (2009) Two types of dopamine neuron distinctly convey positive and negative motivational signals. Nature 459:837–841.

Matsumoto M, Weickert CS, Akil M, Lipska BK, Hyde TM, Herman MM, Kleinman JE, Weinberger DR (2003) Catechol O-methyltransferase mRNA expression in human and rat brain: evidence for a role in cortical neuronal function. Neuroscience 116:127–137.

Mattay VS, Callicott JH, Bertolino A, Heaton I, Frank JA, Coppola R, Berman KF, Goldberg TE, Weinberger DR (2000) Effects of dextroamphetamine on cognitive performance and cortical activation. NeuroImage 12:268–275.

Mattay VS, Goldberg TE, Fera F, Hariri AR, Tessitore A, Egan MF, Kolachana B, Callicott JH, Weinberger DR (2003) Catechol O-methyltransferase Val158Met genotype and individual variation in the brain response to amphetamine. Proc Natl Acad Sci 100:6186–6191.

Ma WJ, Beck JM, Latham PE, Pouget A (2006) Bayesian inference with probabilistic population codes. Nat Neurosci 9:1432–1438.

Ma WJ, Jazayeri M (2014) Neural coding of uncertainty and probability. Annu Rev Neurosci 37:205–220.

McGaughy J, Ross RS, Eichenbaum H (2008) Noradrenergic, but not cholinergic, deafferentation of prefrontal cortex impairs attentional set-shifting. Neuroscience 153:63–71.

McGinley MJ, Vinck M, Reimer J, Batista-Brito R, Zagha E, Cadwell CR, Tolias AS, Cardin JA, McCormick DA (2015) Waking state: Rapid variations modulate neural and behavioral responses. Neuron 87:1143–1161.

McGuire JT, Nassar MR, Gold JI, Kable JW (2014) Functionally dissociable influences on learning rate in a dynamic environment. Neuron 84:870–881.

Mehta MA, Goodyer IM, Sahakian BJ (2004a) Methylphenidate improves working memory and set-shifting in AD/HD: relationships to baseline memory capacity. J Child Psychol Psychiatry 45:293–305.

Mehta MA, Manes FF, Magnolfi G, Sahakian BJ, Robbins TW (2004b) Impaired set-shifting and dissociable effects on tests of spatial working memory following the dopamine D2 receptor antagonist sulpiride in human volunteers. Psychopharmacology 176:331–342.

Mehta MA, Owen AM, Sahakian BJ, Mavaddat N, Pickard JD, Robbins TW (2000) Methylphenidate Enhances working memory by modulating discrete frontal and parietal lobe regions in the human brain. J Neurosci 20:RC65.

Meintzschel F, Ziemann U (2006) Modification of practice-dependent plasticity in human motor cortex by neuromodulators. Cereb Cortex 16:1106–1115.

Meyer-Lindenberg A, Kohn PD, Kolachana B, Kippenhan S, McInerney-Leo A, Nussbaum R, Weinberger DR, Berman KF (2005) Midbrain dopamine and prefrontal function in humans: interaction and modulation by COMT genotype. Nat Neurosci 8:594–596.

Meyer-Lindenberg A, Nichols T, Callicott JH, Ding J, Kolachana B, Buckholtz J, Mattay VS, Egan M, Weinberger DR (2006) Impact of complex genetic variation in COMT on human brain function. Mol Psychiatry 11:867–877, 797.

Meyniel F, Sigman M, Mainen ZF (2015) Confidence as Bayesian Probability: From Neural Origins to Behavior. Neuron 88:78–92.

Mill J, Asherson P, Browes C, D'Souza U, Craig I (2002) Expression of the dopamine transporter gene is regulated by the 3′ UTR VNTR: evidence from brain and lymphocytes using quantitative RT-PCR. Am J Med Genet 114:975–979.

Mirenowicz J, Schultz W (1994) Importance of unpredictability for reward responses in primate dopamine neurons. J Neurophysiol 72:1024–1027.

Moghaddam B, Bunney BS (1990) Acute effects of typical and atypical antipsychotic drugs on the release of dopamine from prefrontal cortex, nucleus accumbens, and striatum of the rat: an in vivo microdialysis study. J Neurochem 54:1755–1760.

Montague PR, Dayan P, Sejnowski TJ (1996) A framework for mesencephalic dopamine systems based on predictive Hebbian learning. J Neurosci 16:1936–1947.

Montague PR, Hyman SE, Cohen JD (2004) Computational roles for dopamine in behavioural control. Nature 431:760–767.

Montana G, Pritchard JK (2004) Statistical tests for admixture mapping with case-control and cases-only data. Am J Hum Genet 75:771–789.

Moore RY, Bloom FE (1979) Central catecholamine neuron systems: anatomy and physiology of the norepinephrine and epinephrine systems. Annu Rev Neurosci 2:113–168.

Moran RJ, Campo P, Symmonds M, Stephan KE, Dolan RJ, Friston KJ (2013) Free energy, precision and learning: the role of cholinergic neuromodulation. J Neurosci 33:8227–8236.

Murphy PR, O'Connell RG, O'Sullivan M, Robertson IH, Balsters JH (2014) Pupil diameter covaries with BOLD activity in human locus coeruleus. Hum Brain Mapp 35:4140–4154.

Murray GK, Corlett PR, Clark L, Pessiglione M, Blackwell AD, Honey G, Jones PB, Bullmore ET, Robbins TW, Fletcher PC (2008) Substantia nigra/ventral tegmental reward prediction error disruption in psychosis. Mol Psychiatry 13:267–276.

Näätänen R (1970) The diminishing time-uncertainty with the lapse of time after the warning signal in reaction-time experiments with varying fore-periods. Acta Psychol 34:399–419.

Nackley AG, Shabalina SA, Tchivileva IE, Satterfield K, Korchynskyi O, Makarov SS, Maixner W, Diatchenko L (2006) Human catechol-O-methyltransferase haplotypes modulate protein expression by altering mRNA secondary structure. Science 314:1930–1933.

Nakahara H, Itoh H, Kawagoe R, Takikawa Y, Hikosaka O (2004) Dopamine neurons can represent context-dependent prediction error. Neuron 41:269–280.

Nassar MR, Rumsey KM, Wilson RC, Parikh K, Heasly B, Gold JI (2012) Rational regulation of learning dynamics by pupil-linked arousal systems. Nat Neurosci 15:1040–1046.

Nassar MR, Wilson RC, Heasly B, Gold JI (2010) An approximately Bayesian delta-rule model explains the dynamics of belief updating in a changing environment. J Neurosci 30:12366–12378.

Newman LA, Darling J, McGaughy J (2008) Atomoxetine reverses attentional deficits produced by noradrenergic deafferentation of medial prefrontal cortex. Psychopharmacology 200:39–50.

Nieuwenhuis S, de Geus EJ, Aston-Jones G (2010) The anatomical and functional relationship between the P3 and autonomic components of the orienting response. Psychophysiology. 48:162-175

Noble EP (2003) D2 dopamine receptor gene in psychiatric and neurologic disorders and its phenotypes. Am J Med Genet B Neuropsychiatr Genet 116:103–125.

O'Doherty J, Dayan P, Schultz J, Deichmann R, Friston K, Dolan RJ (2004) Dissociable roles of ventral and dorsal striatum in instrumental conditioning. Science 304:452–454.

O'Doherty JP, Dayan P, Friston K, Critchley H, Dolan RJ (2003) Temporal difference models and reward-related learning in the human brain. Neuron 38:329–337.

Oh SY, Kim YK (2016) Association of norepinephrine transporter gene polymorphisms in attention-deficit/hyperactivity disorder in Korean population. Prog Neuropsychopharmacol Biol Psychiatry. Epub ahead of print.

O'Neill J, Siembieda DW, Crawford KC, Halgren E, Fisher A, Fitten LJ (2003) Reduction in distractibility with AF102B and THA in the macaque. Pharmacol Biochem Behav 76:301–306.

Orbán G, Fiser J, Aslin RN, Lengyel M (2008) Bayesian learning of visual chunks by human observers. Proc Natl Acad Sci 105:2745–2750.

O'Reilly JX (2013) Making predictions in a changing world-inference, uncertainty, and learning. Front Neurosci 7:105.

O'Reilly JX, Jbabdi S, Behrens TEJ (2012) How can a Bayesian approach inform neuroscience? Eur J Neurosci 35:1169–1179.

Owens MJ, Krulewicz S, Simon JS, Sheehan DV, Thase ME, Carpenter DJ, Plott SJ, Nemeroff CB (2008) Estimates of serotonin and norepinephrine transporter inhibition in depressed patients treated with paroxetine or venlafaxine. Neuropsychopharmacology 33:3201–3212.

Palacios JM, Camps M, Cortés R, Probst A (1988) Mapping dopamine receptors in the human brain. J Neural Transm Suppl 27:227–235.

Palmatier MA, Kang AM, Kidd KK (1999) Global variation in the frequencies of functionally different catechol-O-methyltransferase alleles. Biol Psychiatry 46:557–567.

Parasuraman R, Greenwood PM, Haxby JV, Grady CL (1992) Visuospatial attention in dementia of the Alzheimer type. Brain 115:711–733.

Parkinson JA, Olmstead MC, Burns LH, Robbins TW, Everitt BJ (1999) Dissociation in effects of lesions of the nucleus accumbens core and shell on appetitive pavlovian approach behavior and the potentiation of conditioned reinforcement and locomotor activity by D-amphetamine. J Neurosci 19:2401–2411.

Patton JH, Stanford MS, Barratt ES (1995) Factor structure of the Barratt impulsiveness scale. J Clin Psychol:768–774.

Payzan-LeNestour E, Bossaerts P (2011) Risk, unexpected uncertainty, and estimation uncertainty: Bayesian learning in unstable settings. PLOS Comput Biol 7:e1001048.

Payzan-LeNestour E, Dunne S, Bossaerts P, O'Doherty JP (2013) The neural representation of unexpected uncertainty during value-based decision making. Neuron 79:191–201.

Pearson-Fuhrhop KM, Minton B, Acevedo D, Shahbaba B, Cramer SC (2013) Genetic Variation in the human brain dopamine system influences motor learning and its modulation by L-Dopa. PLOS ONE 8:e61197.

Pessiglione M, Seymour B, Flandin G, Dolan RJ, Frith CD (2006) Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. Nature 442:1042–1045.

Peterson DA, Elliott C, Song DD, Makeig S, Sejnowski TJ, Poizner H (2009) Probabilistic reversal learning is impaired in Parkinson's disease. Neuroscience 163:1092–1101.

Peterson GB (2004) A day of great illumination: B. F. Skinner's discovery of shaping. J Exp Anal Behav 82:317–328.

Phillips JM, McAlonan K, Robb WGK, Brown VJ (2000a) Cholinergic neurotransmission influences covert orientation of visuospatial attention in the rat. Psychopharmacology 150:112.

Phillips MA, Szabadi E, Bradshaw CM (2000b) Comparison of the effects of clonidine and yohimbine on spontaneous pupillary fluctuations in healthy human volunteers. Psychopharmacology 150:85–89.

Phillips MA, Szabadi E, Bradshaw CM (2000c) Comparison of the effects of clonidine and yohimbine on pupillary diameter at different illumination levels. Br J Clin Pharmacol 50:65–68.

Polack P-O, Friedman J, Golshani P (2013) Cellular mechanisms of brain state-dependent gain modulation in visual cortex. Nat Neurosci 16:1331–1339.

Posner MI (1980) Orienting of attention. Q J Exp Psychol 32:3–25.

Pouget A, Beck JM, Ma WJ, Latham PE (2013) Probabilistic brains: knowns and unknowns. Nat Neurosci 16:1170–1178.

Pouget A, Dayan P, Zemel RS (2003) Inference and computation with population codes. Annu Rev Neurosci 26:381–410.

Pouget A, Drugowitsch J, Kepecs A (2016) Confidence and certainty: distinct probabilistic quantities for different goals. Nat Neurosci 19:366–374.

Prendergast MA, Jackson WJ, Terry Jr AV, Decker MW, Arneric SP, Buccafusco JJ (1998) Central nicotinic receptor agonists ABT-418, ABT-089, and (–)-nicotine reduce distractibility in adult monkeys. Psychopharmacology 136:50–58.

Preston AR, Shrager Y, Dudukovic NM, Gabrieli JDE (2004) Hippocampal contribution to the novel use of relational information in declarative memory. Hippocampus 14:148–152.

Preuschoff K, Bossaerts P (2007) Adding prediction risk to the theory of reward learning. Ann N Y Acad Sci 1104:135–146.

Preuschoff K, 't Hart BM, Einhäuser W (2011) Pupil dilation signals surprise: Evidence for noradrenaline's role in decision making. Front Neurosci 5:1–12.

Privitera CM, Renninger LW, Carney T, Klein S, Aguilar M (2010) Pupil dilation during visual target detection. J Vis 10:3.

Purcell SM, Wray NR, Stone JL, Visscher PM, O'Donovan MC, Sullivan PF, Sklar P (2009) Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. Nature 460:748–752.

Purves D, Augustine GJ, Fitzpatrick D, Hall WC, LaMantia A-S, White LE (2011) Neuroscience. Sinauer Associates.

Pycock CJ, Kerwin RW, Carter CJ (1980) Effect of lesion of cortical dopamine terminals on subcortical dopamine receptors in rats. Nature 286:74–76.

Rabbitt PM (1966) Errors and error correction in choice-response tasks. J Exp Psychol 71:264.

Rajkowski J, Kubiak P, Aston-Jones G (1993) Correlations between locus coeruleus (LC) neural activity, pupil diameter and behavior in monkey support a role of LC in attention. In: Society for Neuroscience Abstracts, pp 974.

Rajkowski J, Majczynski H, Clayton E, Aston-Jones G (2004) Activation of monkey locus coeruleus neurons varies with difficulty and performance in a target detection task. J Neurophysiol 92:361–371.

Redgrave P, Prescott TJ, Gurney K (1999) Is the short-latency dopamine response too short to signal reward error? Trends Neurosci 22:146–151.

Reimer J, McGinley MJ, Liu Y, Rodenkirch C, Wang Q, McCormick DA, Tolias AS (2016) Pupil fluctuations track rapid changes in adrenergic and cholinergic activity in cortex. Nat Commun 7:13289.

Requin J, Granjon M (1969) The effect of conditional probability of the response signal on the simple reaction time. Acta Psychol 31:129–144.

Rescorla RA, Wagner AR (1972) A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In: Classical conditioning: Current research and theory.

Richfield EK, Penney JB, Young AB (1989) Anatomical and affinity state comparisons between dopamine D1 and D2 receptors in the rat central nervous system. Neuroscience 30:767–777.

Rigoux L, Stephan KE, Friston KJ, Daunizeau J (2014) Bayesian model selection for group studies—revisited. Neuroimage 84:971–985.

Ritchie T, Noble EP (2003) Association of seven polymorphisms of the D2 dopamine receptor gene with brain receptor-binding characteristics. Neurochem Res 28:73–82.

Rodriguez-Murillo DL, Salem DRM (2013) Insertion/Deletion Polymorphism. In: Encyclopedia of Behavioral Medicine. Springer New York.

Roesch-Ely D, Scheffel H, Weiland S, Schwaninger M, Hundemer H-P, Kolter T, Weisbrod M (2005) Differential dopaminergic modulation of executive control in healthy subjects. Psychopharmacology 178:420–430.

Rushworth MF, Behrens TE (2008) Choice, uncertainty and value in prefrontal and cingulate cortex. Nat Neurosci 11:389–397.

Rutledge RB, Skandali N, Dayan P, Dolan RJ (2014) A computational and neural model of momentary subjective well-being. Proc Natl Acad Sci 111:12252–12257.

Rutledge RB, Skandali N, Dayan P, Dolan RJ (2015) Dopaminergic modulation of decision making and subjective well-being. J Neurosci 35:9811–9822.

Sachidanandam R et al. (2001) A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. Nature 409:928–933.

Samuels R, Szabadi E (2008) Functional neuroanatomy of the noradrenergic locus coeruleus: its roles in the regulation of arousal and autonomic function part II: physiological and pharmacological manipulations and pathological alterations of locus coeruleus activity in humans. Curr Neuropharmacol 6:254-285.

Sanger TD (1996) Probability density estimation for the interpretation of neural population codes. J Neurophysiol 76:2790–2793.

Santangelo G, Piscopo F, Barone P, Vitale C (2017) Personality in Parkinson's disease: Clinical, behavioural and cognitive correlates. J Neurol Sci 374:17–25.

Sara SJ (2009) The locus coeruleus and noradrenergic modulation of cognition. Nat Rev Neurosci 10:211–223.

Sara SJ, Segal M (1991) Plasticity of sensory responses of locus coeruleus neurons in the behaving rat: implications for cognition. Prog Brain Res 88:571–585.

Sara SJ, Vankov A, Hervé A (1994) Locus coeruleus-evoked responses in behaving rats: a clue to the role of noradrenaline in memory. Brain Res Bull 35:457–465.

Sarter M, Hasselmo ME, Bruno JP, Givens B (2005) Unraveling the attentional functions of cortical cholinergic inputs: interactions between signal-driven and cognitive modulation of signal detection. Brain Res Rev 48:98–111.

Satterthwaite TD, Green L, Myerson J, Parker J, Ramaratnam M, Buckner RL (2007) Dissociable but inter-related systems of cognitive control and reward during decision making: Evidence from pupillometry and event-related fMRI. NeuroImage 37:1017–1031.

Scacchi R, Gambina G, Moretto G, Corbo RM (2009) Variability of AChE, BChE, and ChAT genes in the late-onset form of Alzheimer's disease and relationships with

response to treatment with Donepezil and Rivastigmine. Am J Med Genet B Neuropsychiatr Genet 150:502–507.

Schmitz Y, Benoit-Marand M, Gonon F, Sulzer D (2003) Presynaptic regulation of dopaminergic neurotransmission. J Neurochem 87:273–289.

Schoemaker H, Claustre Y, Fage D, Rouquier L, Chergui K, Curet O, Oblin A, Gonon F, Carter C, Benavides J, Scatton B (1997) Neurochemical characteristics of amisulpride, an atypical dopamine D2/D3 receptor antagonist with both presynaptic and limbic selectivity. J Pharmacol Exp Ther 280:83–97.

Schultz W, Dayan P, Montague PR (1997) A neural substrate of prediction and reward. Science 275:1593–1599.

Schultz W, Dickinson A (2000) Neuronal coding of prediction errors. Annu Rev Neurosci 23:473–500.

Schwinn DA, Johnston GI, Page SO, Mosley MJ, Wilson KH, Worman NP, Campbell S, Fidock MD, Furness LM, Parry-Smith DJ (1995) Cloning and pharmacological characterization of human alpha-1 adrenergic receptors: sequence corrections and direct comparison with other species homologues. J Pharmacol Exp Ther 272:134–142.

Sekine M, Arakawa R, Ito H, Okumura M, Sasaki T, Takahashi H, Takano H, Okubo Y, Halldin C, Suhara T (2010) Norepinephrine transporter occupancy by antidepressant in human brain using positron emission tomography with (S,S)-[18F]FMeNER-D2. Psychopharmacology 210:331–336.

Servan-Schreiber D, Printz H, Cohen JD (1990) A Network Model of Catecholamine Effects: Gain, Signal-to-Noise Ratio, and Behavior. Science 249:892–895.

Sesack SR, Hawrylak VA, Matus C, Guido MA, Levey AI (1998) Dopamine axon varicosities in the prelimbic division of the rat prefrontal cortex exhibit sparse immunoreactivity for the dopamine transporter. J Neurosci 18:2697–2708.

Seu E, Lang A, Rivera RJ, Jentsch JD (2009) Inhibition of the norepinephrine transporter improves behavioral flexibility in rats and monkeys. Psychopharmacology 202:505–519.

Seymour B, Daw N, Dayan P, Singer T, Dolan R (2007) Differential Encoding of Losses and Gains in the Human Striatum. J Neurosci 27:4826–4831.

Shibata E, Sasaki M, Tohyama K, Kanbara Y, Otsuka K, Ehara S, Sakai A (2006) Age-related changes in locus ceruleus on neuromelanin magnetic resonance imaging at 3 Tesla. Magn Reson Med Sci 5:197–200.

Shi J et al. (2009) Common variants on chromosome 6p22.1 are associated with schizophrenia. Nature 460:753–757.

Shohamy D, Wagner AD (2008) Integrating Memories in the Human Brain: Hippocampal-Midbrain Encoding of Overlapping Events. Neuron 60:378–389.

Siegle GJ, Steinhauer SR, Stenger VA, Konecky R, Carter CS (2003) Use of concurrent pupil dilation assessment to inform interpretation and analysis of fMRI data. Neuroimage 20:114–124.

Silvetti M, Seurinck R, Verguts T (2013) Value and prediction error estimation account for volatility effects in ACC: a model-based fMRI study. Cortex 49:1627–1635.

Slifstein M, Kolachana B, Simpson EH, Tabares P, Cheng B, Duvall M, Frankle WG, Weinberger DR, Laruelle M, Abi-Dargham A (2008) COMT genotype predicts cortical-limbic D1 receptor availability measured with [11C]NNC112 and PET. Mol Psychiatry 13:821–827.

Smith A, Li M, Becker S, Kapur S (2006) Dopamine, prediction error and associative learning: a model-based account. Netw Comput Neural Syst 17:61–84.

Smith AP, Wilson SJ, Glue P, Nutt DJ (1992) The effects and after effects of the α2-adrenoceptor antagonist idazoxan on mood, memory and attention in normal volunteers. J Psychopharmacol 6:376–381.

Solís-Ortiz S, Pérez-Luque E, Morado-Crespo L, Gutiérrez-Muñoz M (2010) Executive functions and selective attention are favored in middle-aged healthy women carriers of the Val/Val genotype of the catechol-o-methyltransferase gene: a behavioral genetic study. Behav Brain Funct 6:67.

Soranzo A, Grassi M (2014) PSYCHOACOUSTICS: a comprehensive MATLAB toolbox for auditory testing. Front Psychol 5:712.

Spencer TJ, Biederman J, Faraone SV, Madras BK, Bonab AA, Dougherty DD, Batchelder H, Clarke A, Fischman AJ (2013) Functional genomics of attention-deficit/hyperactivity disorder (ADHD) risk alleles on dopamine transporter binding in ADHD and healthy control subjects. Biol Psychiatry 74:84–89.

Starke K, Göthert M, Kilbinger H (1989) Modulation of neurotransmitter release by presynaptic autoreceptors. Physiol Rev 69:864–989.

Stefansson H et al. (2009) Common variants conferring risk of schizophrenia. Nature 460:744–747.

Stelzel C, Basten U, Montag C, Reuter M, Fiebach CJ (2010) Frontostriatal involvement in task switching depends on genetic differences in d2 receptor density. J Neurosci 30:14205–14212.

Stelzel C, Fiebach CJ, Cools R, Tafazoli S, D'Esposito M (2013) Dissociable fronto-striatal effects of dopamine D2 receptor stimulation on cognitive versus motor flexibility. Cortex 49:2799–2811.

Stephan KE, Penny WD, Daunizeau J, Moran RJ, Friston KJ (2009) Bayesian model selection for group studies. Neuroimage 46:1004–1017.

Stephenson RP (1997) A Modification of Receptor Theory. Br J Pharmacol 120:106–120.

Sterpenich V, D'Argembeau A, Desseilles M, Balteau E, Albouy G, Vandewalle G, Degueldre C, Luxen A, Collette F, Maquet P (2006) The locus ceruleus is involved in the successful retrieval of emotional memories in humans. J Neurosci 26:7416–7423.

Steyvers M, Lee MD, Wagenmakers E-J (2009) A Bayesian analysis of human decision-making on bandit problems. J Math Psychol 53:168–179.

Stone EA, Lin Y, Ahsan R, Quartermain D (2004) Role of locus coeruleus α1-adrenoceptors in motor activity in rats. Synapse 54:164–172.

Sutton RS (1988) Learning to predict by the methods of temporal differences. Mach Learn 3:9–44.

Sutton RS (1992) Gain adaptation beats least squares. In: Proceedings of the 7th Yale workshop on adaptive and learning systems, pp 161–166.

Sutton RS, Barto AG (1981) Toward a modern theory of adaptive networks: expectation and prediction. Psychol Rev 88:135–170.

Sutton RS, Barto AG (1998) Reinforcement learning: An introduction. MIT Press Cambridge.

Sutton S, Tueting P, Zubin J, John ER (1967) Information delivery and the sensory evoked potential. Science 155:1436–1439.

Swanson CJ, Perry KW, Koch-Krueger S, Katner J, Svensson KA, Bymaster FP (2006) Effect of the attention deficit/hyperactivity disorder drug atomoxetine on extracellular concentrations of norepinephrine and dopamine in several brain regions of the rat. Neuropharmacology 50:755–760.

Tait DS, Brown VJ, Farovik A, Theobald DE, Dalley JW, Robbins TW (2007) Lesions of the dorsal noradrenergic bundle impair attentional set-shifting in the rat. Eur J Neurosci 25:3719–3724.

Takahashi YK, Roesch MR, Wilson RC, Toreson K, O'Donnell P, Niv Y, Schoenbaum G (2011) Expectancy-related changes in firing of dopamine neurons depend on orbitofrontal cortex. Nat Neurosci 14:1590–1597.

Tan H-Y, Callicott JH, Weinberger DR (2007a) Dysfunctional and compensatory prefrontal cortical systems, genes and the pathogenesis of schizophrenia. Cereb Cortex 17 Suppl 1:i171–i181.

Tan H-Y, Chen Q, Goldberg TE, Mattay VS, Meyer-Lindenberg A, Weinberger DR, Callicott JH (2007b) Catechol-O-methyltransferase Val158Met modulation of prefrontal-parietal-striatal brain systems during arithmetic and temporal transformations in working memory. J Neurosci Off J Soc Neurosci 27:13393–13401.

Tenenbaum JB, Griffiths TL, Kemp C (2006) Theory-based Bayesian models of inductive learning and reasoning. Trends Cogn Sci 10:309–318.

Terry AV, Risbrough VB, Buccafusco JJ, Menzaghi F (2002) Effects of (±)-4-{[2-(1-methyl-2-pyrrolidinyl)ethyl]thio}phenol hydrochloride (SIB-1553A), a selective ligand for nicotinic acetylcholine receptors, in tests of visual attention and distractibility in rats and monkeys. J Pharmacol Exp Ther 301:284–292.

Thiel CM, Bentley P, Dolan RJ (2002) Effects of cholinergic enhancement on conditioning-related responses in human auditory cortex. Eur J Neurosci 16:2199–2206.

Thiel CM, Henson RNA, Morris JS, Friston KJ, Dolan RJ (2001) Pharmacological modulation of behavioral and neuronal correlates of repetition priming. J Neurosci 21:6846–6852.

Thoma P, Wiebel B, Daum I (2007) Response inhibition and cognitive flexibility in schizophrenia with and without comorbid substance use disorder. Schizophr Res 92:168–180.

Thompson J, Thomas N, Singleton A, Piggott M, Lloyd S, Perry EK, Morris CM, Perry RH, Ferrier IN, Court JA (1997) D2 dopamine receptor gene (DRD2) Taq1 A polymorphism: reduced dopamine D2 receptor binding in the human striatum associated with the A1 allele. Pharmacogenetics 7:479–484.

Tomassini A, Ruge D, Galea JM, Penny W, Bestmann S (2015) The role of dopamine in temporal uncertainty. J Cogn Neurosci 28:96–110.

Tully K, Bolshakov VY (2010) Emotional enhancement of memory: how norepinephrine enables synaptic plasticity. Mol Brain 3:1.

Tunbridge EM (2010) The catechol-O-methyltransferase gene: its regulation and polymorphisms. Int Rev Neurobiol 95:7–27.

Tunbridge EM, Bannerman DM, Sharp T, Harrison PJ (2004) Catechol-o-methyltransferase inhibition improves set-shifting performance and elevates stimulated dopamine release in the rat prefrontal cortex. J Neurosci 24:5331–5335.

Tunbridge EM, Harrison PJ, Weinberger DR (2006) Catechol-o-methyltransferase, cognition, and psychosis: Val158Met and beyond. Biol Psychiatry 60:141–151.

Ullsperger M (2010) Genetic association studies of performance monitoring and learning from feedback: the role of dopamine and serotonin. Neurosci Biobehav Rev 34:649–659.

Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, Rozen SG (2012) Primer3 - new capabilities and interfaces. Nucleic Acids Res 40:e115–e115.

Urai AE, Braun A, Donner TH (2017) Pupil-linked arousal is driven by decision uncertainty and alters serial choice bias. Nat Commun 8:14637.

Usher M, Cohen JD, Servan-Schreiber D, Rajkowski J, Aston-Jones G (1999) The role of locus coeruleus in the regulation of cognitive performance. Science 283:549–554.

van de Giessen EM, de Win MM, Tanck MW, van den Brink W, Baas F, Booij J (2009) Striatal dopamine transporter availability associated with polymorphisms in the dopamine transporter gene SLC6A3. J Nucl Med 50:45–52.

van den Brink RL, Murphy PR, Nieuwenhuis S (2016) Pupil diameter tracks lapses of attention. PloS One 11:e0165274.

van der Schaaf ME, van Schouwenburg MR, Geurts DE, Schellekens AF, Buitelaar JK, Verkes RJ, Cools R (2014) Establishing the dopamine dependency of human striatal signals during reward and punishment reversal learning. Cereb Cortex 24:633–642.

van Dyck CH, Malison RT, Jacobsen LK, Seibyl JP, Staley JK, Laruelle M, Baldwin RM, Innis RB, Gelernter J (2005) Increased dopamine transporter availability associated with the 9-repeat allele of the SLC6A3 gene. J Nucl Med 46:745–751.

van Holstein M, Aarts E, van der Schaaf ME, Geurts DE, Verkes RJ, Franke B, van Schouwenburg MR, Cools R (2011) Human cognitive flexibility depends on dopamine D2 receptor signaling. Psychopharmacology 218:567–578.

Vankov A, Hervé-Minvielle A, Sara SJ (1995) Response to novelty and its rapid habituation in locus coeruleus neurons of the freely exploring rat. Eur J Neurosci 7:1180–1187.

VanNess SH, Owens MJ, Kilts CD (2005) The variable number of tandem repeats element in DAT1 regulates in vitro dopamine transporter density. BMC Genet 6:55.

Varazzani C, San-Galli A, Gilardeau S, Bouret S (2015) Noradrenaline and dopamine neurons in the reward/effort trade-off: A direct electrophysiological comparison in behaving monkeys. J Neurosci 35:7866–7877.

Vijayraghavan S, Wang M, Birnbaum SG, Williams GV, Arnsten AFT (2007) Inverted-U dopamine D1 receptor actions on prefrontal neurons engaged in working memory. Nat Neurosci 10:376–384.

Volkmann J (1934) The relation of the time of judgment to the certainty of judgment. Psychol Bull 31:672–673.

Volpicelli LA, Levey AI (2004) Muscarinic acetylcholine receptor subtypes in cerebral cortex and hippocampus. Prog Brain Res 145:59-66.

Vossel S, Bauer M, Mathys C, Adams RA, Dolan RJ, Stephan KE, Friston KJ (2014a) Cholinergic stimulation enhances Bayesian belief updating in the deployment of spatial attention. J Neurosci 34:15735–15742.

Vossel S, Mathys C, Daunizeau J, Bauer M, Driver J, Friston KJ, Stephan KE (2014b) Spatial attention, precision, and Bayesian inference: A study of saccadic response speed. Cereb Cortex 24:1436–1450.

Vossel S, Mathys C, Stephan KE, Friston KJ (2015) Cortical coupling reflects Bayesian belief updating in the deployment of spatial attention. J Neurosci 35:11532–11542.

Voytko ML, Olton DS, Richardson RT, Gorman LK, Tobin JR, Price DL (1994) Basal forebrain lesions in monkeys disrupt attention but not learning and memory [published erratum appears in J Neurosci 1995 Mar;15(3): following table of contents]. J Neurosci 14:167–186.

Wang C-A, Boehnke SE, White BJ, Munoz DP (2012) Microstimulation of the monkey superior colliculus induces pupil dilation without evoking saccades. J Neurosci 32:3629–3636.

Warren CM, Eldar E, Brink RL van den, Tona K-D, Wee NJ van der, Giltay EJ, Noorden MS van, Bosch JA, Wilson RC, Cohen JD, Nieuwenhuis S (2016) Catecholamine-mediated increases in gain enhance the precision of cortical representations. J Neurosci 36:5699–5708.

Wichmann T, DeLong MR (1996) Functional and pathophysiological models of the basal ganglia. Curr Opin Neurobiol 6:751–758.

Wickens JR, Reynolds JNJ, Hyland BI (2003) Neural mechanisms of reward-related motor learning. Curr Opin Neurobiol 13:685–690.

Wierda SM, van Rijn H, Taatgen NA, Martens S (2012) Pupil dilation deconvolution reveals the dynamics of attention at high temporal resolution. Proc Natl Acad Sci 109:8456–8460.

Williams GV, Goldman-Rakic PS (1995) Modulation of memory fields by dopamine D1 receptors in prefrontal cortex. Nature 376:572–575.

Wilson RC, Nassar MR, Gold JI (2010) Bayesian online learning of the hazard rate in change-point problems. Neural Comput 22:2452–2476.

Wilson RC, Nassar MR, Gold JI (2013) A mixture of delta-rules approximation to Bayesian inference in change-point problems. PLOS Comput Biol 9:e1003150.

Winterer G, Goldman D (2003) Genetics of human prefrontal function. Brain Res Rev 43:134–163.

Winterer G, Weinberger DR (2004) Genes, dopamine and cortical signal-to-noise ratio in schizophrenia. Trends Neurosci 27:683–690.

Wise RA (2004) Dopamine, learning and motivation. Nat Rev Neurosci 5:483–494.

Witte EA, Davidson MC, Marrocco RT (1997) Effects of altering brain cholinergic activity on covert orienting of attention: comparison of monkey and human performance. Psychopharmacology 132:324-334.

Witte EA, Marrocco RT (1997) Alteration of brain noradrenergic activity in rhesus monkeys affects the alerting component of covert orienting. Psychopharmacology (Berl) 132:315–323.

Wong DF, Wagner HN, Tune LE, Dannals RF, Pearlson GD, Links JM, Tamminga CA, Broussolle EP, Ravert HT, Wilson AA, others (1986) Positron emission tomography reveals elevated D2 dopamine receptors in drug-naive schizophrenics. Science 234:1558–1563.

Wong EHF, Sonders MS, Amara SG, Tinholt PM, Piercey MFP, Hoffmann WP, Hyslop DK, Franklin S, Porsolt RD, Bonsignori A, Carfagna N, McArthur RA (2000) Reboxetine: a pharmacologically potent, selective, and specific norepinephrine reuptake inhibitor. Biol Psychiatry 47:818–829.

Wu Q, Reith MEA, Walker QD, Kuhn CM, Carroll FI, Garris PA (2002) Concurrent autoreceptor-mediated control of dopamine release and uptake during neurotransmission: an in vivo voltammetric study. J Neurosci 22:6272–6281.

Xu F, Tenenbaum JB (2007) Sensitivity to sampling in Bayesian word learning. Dev Sci 10:288–297.

Yang H, Wang K (2015) Genomic variant annotation and prioritization with ANNOVAR and wANNOVAR. Nat Protoc 10:1556–1566.

Yang T, Shadlen MN (2007) Probabilistic reasoning by neurons. Nature 447:1075–1080.

Yavich L, Forsberg MM, Karayiorgou M, Gogos JA, Männistö PT (2007) Site-specific role of catechol-O-methyltransferase in dopamine overflow within prefrontal cortex and dorsal striatum. J Neurosci 27:10196–10209.

Yeung N, Botvinick MM, Cohen JD (2004) The neural basis of error detection: conflict monitoring and the error-related negativity. Psychol Rev 111:931.

Yu A, Dayan P (2003) Expected and unexpected uncertainty: ACh and NE in the neocortex. Adv Neural Inf Process Syst:173–180.

Yu AJ (2007) Adaptive Behavior: Humans Act as Bayesian Learners. Curr Biol 17:R977–R980.

Yu AJ (2012) Change is in the eye of the beholder. Nat Neurosci 15:933–935.

Yu AJ, Dayan P (2002) Acetylcholine in cortical inference. Neural Netw 15:719–730.

Yu AJ, Dayan P (2005) Uncertainty, neuromodulation, and attention. Neuron 46:681–692.

Yu AJ, Dayan P, Cohen JD (2009) Dynamics of Attentional Selection under Conflict: Toward a Rational Bayesian Account. J Exp Psychol Hum Percept Perform 35:700–717.

Yuille A, Kersten D (2006) Vision as Bayesian inference: analysis by synthesis? Trends Cogn Sci 10:301–308.

Zaghloul KA, Blanco JA, Weidemann CT, McGill K, Jaggi JL, Baltuch GH, Kahana MJ (2009) Human substantia nigra neurons encode unexpected financial rewards. Science 323:1496–1499.

Zhao X, Huang Y, Ma H, Jin Q, Wang Y, Zhu G (2013) Association between major depressive disorder and the norepinephrine transporter polymorphisms T-182C and G1287A: a meta-analysis. J Affect Disord 150:23–28.

Zhu J, Taniguchi T, Takauji R, Suzuki F, Tanaka T, Muramatsu I (2000) Inverse agonism and neutral antagonism at a constitutively active alpha-1a adrenoceptor. Br J Pharmacol 131:546–552.

Ziemann U, Tergau F, Bruns D, Baudewig J, Paulus W (1997) Changes in human motor cortex excitability induced by dopaminergic and anti-dopaminergic drugs. Electroencephalogr Clin Neurophysiol Mot Control 105:430–437.

Zilles K, Qü M, Schleicher A (1993) Regional distribution and heterogeneity of alpha-adrenoceptors in the rat and human central nervous system. J Hirnforsch 34:123–132.

Zink CF, Pagnoni G, Martin ME, Dhamala M, Berns GS (2003) Human striatal response to salient nonrewarding stimuli. J Neurosci 23:8092–8097.